Optical Design Fundamentals for Infrared Systems

Second Edition

Tutorial Texts Series

- *Optical Design Fundamentals for Infrared Systems, Second Edition.*, Max J. Riedl, Vol. TT48
- *Resolution Enhancement Techniques in Optical Lithography*, Alfred Kwok-Kit Wong, Vol. TT47
- *Copper Interconnect Technology*, Christoph Steinbrüchel and Barry L. Chin, Vol. TT46
- *Optical Design for Visual Systems*, Bruce H. Walker, Vol. TT45
- *Fundamentals of Contamination Control*, Alan C. Tribble, Vol. TT44
- *Evolutionary Computation: Principles and Practice for Signal Processing*, David Fogel, Vol. TT43
- *Infrared Optics and Zoom Lenses*, Allen Mann, Vol. TT42
- *Introduction to Adaptive Optics*, Robert K. Tyson, Vol. TT41
- *Fractal and Wavelet Image Compression Techniques*, Stephen Welstead, Vol. TT40
- *Analysis of Sampled Imaging Systems*, R. H. Vollmerhausen and R. G. Driggers, Vol. TT39
- *Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnosis*, Valery Tuchin, Vol. TT38
- *Fundamentos de Electro-Óptica para Ingenieros*, Glenn D. Boreman, translated by Javier Alda, Vol. TT37
- *Infrared Design Examples*, William L. Wolfe, Vol. TT36
- *Sensor and Data Fusion Concepts and Applications, Second Edition*, L. A. Klein, Vol. TT35
- *Practical Applications of Infrared Thermal Sensing and Imaging Equipment, Second Edition*, Herbert Kaplan, Vol. TT34
- *Fundamentals of Machine Vision*, Harley R. Myler, Vol. TT33
- *Design and Mounting of Prisms and Small Mirrors in Optical Instruments*, Paul R. Yoder, Jr., Vol. TT32
- *Basic Electro-Optics for Electrical Engineers*, Glenn D. Boreman, Vol. TT31
- *Optical Engineering Fundamentals*, Bruce H. Walker, Vol. TT30
- *Introduction to Radiometry*, William L. Wolfe, Vol. TT29
- *Lithography Process Control*, Harry J. Levinson, Vol. TT28
- *An Introduction to Interpretation of Graphic Images*, Sergey Ablameyko, Vol. TT27
- *Thermal Infrared Characterization of Ground Targets and Backgrounds*, P. Jacobs, Vol. TT26
- *Introduction to Imaging Spectrometers*, William L. Wolfe, Vol. TT25
- *Introduction to Infrared System Design*, William L. Wolfe, Vol. TT24
- *Introduction to Computer-based Imaging Systems*, D. Sinha, E. R. Dougherty, Vol. TT23
- *Optical Communication Receiver Design*, Stephen B. Alexander, Vol. TT22
- *Mounting Lenses in Optical Instruments*, Paul R. Yoder, Jr., Vol. TT21
- *Optical Design Fundamentals for Infrared Systems*, Max J. Riedl, Vol. TT20
- *An Introduction to Real-Time Imaging*, Edward R. Dougherty, Phillip A. Laplante, Vol. TT19
- *Introduction to Wavefront Sensors*, Joseph M. Geary, Vol. TT18
- *Integration of Lasers and Fiber Optics into Robotic Systems*, J. A. Marszalec, E. A. Marszalec, Vol. TT17
- *An Introduction to Nonlinear Image Processing*, E. R. Dougherty, J. Astola, Vol. TT16
- *Introduction to Optical Testing*, Joseph M. Geary, Vol. TT15
- *Image Formation in Low-Voltage Scanning Electron Microscopy*, L. Reimer, Vol. TT12
- *Diazonaphthoquinone-based Resists*, Ralph Dammel, Vol. TT11
- *Infrared Window and Dome Materials*, Daniel C. Harris, Vol. TT10
- *An Introduction to Morphological Image Processing*, Edward R. Dougherty, Vol. TT9
- *An Introduction to Optics in Computers*, Henri H. Arsenault, Yunlong Sheng, Vol. TT8
- *Digital Image Compression Techniques*, Majid Rabbani, Paul W. Jones, Vol. TT7
- *Aberration Theory Made Simple*, Virendra N. Mahajan, Vol. TT6
- *Single-Frequency Semiconductor Lasers*, Jens Buus, Vol. TT5
- *An Introduction to Biological and Artificial Neural Networks for Pattern Recognition*, Steven K. Rogers, Matthew Kabrisky, Vol. TT4
- *Laser Beam Propagation in the Atmosphere*, Hugo Weichel, Vol. TT3
- *Infrared Fiber Optics*, Paul Klocek, George H. Sigel, Jr., Vol. TT2
- *Spectrally Selective Surfaces for Heating and Cooling Applications*, C. G. Granqvist, Vol. TT1

Optical Design Fundamentals for Infrared Systems

Second Edition

Max J. Riedl

Tutorial Texts in Optical Engineering
Volume TT48

Arthur R. Weeks, Jr., Series Editor
Invivo Research Inc. and University of Central Florida

SPIE PRESS
A Publication of SPIE—The International Society for Optical Engineering
Bellingham, Washington USA

Library of Congress Cataloging-in-Publication Data

Riedl, Max J.
Optical design fundamentals for infrared systems / Max J. Riedl.-- 2nd ed.
p. cm. — (Tutorial texts in optical engineering ; v. TT48)
Includes bibliographical references and index.
ISBN 0-8194-4051-5
1. Optical instruments–Design and construction. 2. Infrared equipment–Design and construction. I. Title. II. Series.

TS514.R53 2001
621.36'2—dc21 00-068758
CIP

Published by

SPIE—The International Society for Optical Engineering
P.O. Box 10
Bellingham, Washington 98227-0010 USA
Phone: (1) 360 676 3290
Fax: (1) 360 647 1445
Email: spie@spie.org
WWW: www.spie.org

Copyright © 2001 The Society of Photo-Optical Instrumentation Engineers

All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means without written permission of the publisher.

Printed in the United States of America.

Introduction to the Series

The Tutorial Texts series was initiated in 1989 as a way to make the material presented in SPIE short courses available to those who couldn't attend and to provide a reference book for those who could. Typically, short course notes are developed with the thought in mind that supporting material will be presented verbally to complement the notes, which are generally written in summary form, highlight key technical topics, and are not intended as stand-alone documents. Additionally, the figures, tables, and other graphically formatted information included with the notes require further explanation given in the instructor's lecture. As stand-alone documents, short course notes do not generally serve the student or reader well.

Many of the Tutorial Texts have thus started as short course notes subsequently expanded into books. The goal of the series is to provide readers with books that cover focused technical interest areas in a tutorial fashion. What separates the books in this series from other technical monographs and textbooks is the way in which the material is presented. Keeping in mind the tutorial nature of the series, many of the topics presented in these texts are followed by detailed examples that further explain the concepts presented. Many pictures and illustrations are included with each text, and where appropriate tabular reference data are also included.

To date, the texts published in this series have encompassed a wide range of topics, from geometrical optics to optical detectors to image processing. Each proposal is evaluated to determine the relevance of the proposed topic. This initial reviewing process has been very helpful to authors in identifying, early in the writing process, the need for additional material or other changes in approach that serve to strengthen the text. Once a manuscript is completed, it is peer reviewed to ensure that chapters communicate accurately the essential ingredients of the processes and technologies under discussion.

During the past nine years, my predecessor, Donald C. O'Shea, has done an excellent job in building the Tutorial Texts series, which now numbers nearly forty books. It has expanded to include not only texts developed by short course instructors but also those written by other topic experts. It is my goal to maintain the style and quality of books in the series, and to further expand the topic areas to include emerging as well as mature subjects in optics, photonics, and imaging.

Arthur R. Weeks, Jr.
Invivo Research Inc. and University of Central Florida

Once more, to Hermine, Renee, Jim, Bryan, and Stephanie.

CONTENTS

Preface xv
Historical Remarks xvii

Chapter 1. Radiometric Considerations

1.1 Introduction 1
1.2 Basic Optical Relations 1
1.3 Signal-to-Noise Ratio 2
1.4 Extended Simplified Radiometric Performance Equation 4
1.5 Thermal Radiation Laws 4
1.5.1 Blackbody radiation and Planck's law 5
1.5.2 The Stefan-Boltzmann law 6
1.5.3 Wien's displacement law 7
1.5.4 Kirchhoff's law and emissivity 7
1.6 Transmission through the Atmosphere 8
1.6.1 Permanent constituents of dry atmosphere 9
1.6.2 Variable constituents 9
1.6.3 Approximation (assumption) 9
1.6.4 Precipitable water (definition) 10
1.6.5 Humidity 11
1.6.6 Precipitable water (calculation) 11
1.6.7 Atmospheric transmission (calculation) 12
1.6.8 Computer models 13
1.7 Typical IR Detectors 14
1.7.1 Thermal detectors 14
1.7.2 Photon or quantum detectors 14
1.7.3 Photoconductive detectors 14
1.7.4 Specific detectivity and noise equivalent bandwidth 15
1.7.5 Detector configurations 16
References 17

Chapter 2. Basic Optics

2.1 Introduction 19
2.2 Snell's Law and the Prism 20
2.3 The Transition from a Prism to a Lens 20
2.4 Image Formation 21
2.5 Object–Image Relations 23
2.6 Stops, Pupils, and Windows 24
2.7 Throughput 26
2.8 Energy Transfer 28
2.8.1 Signal-to-noise calculations 28

2.9 Differential Changes 29
2.10 Optical Gain 30
2.10.1 Immersion lenses 30
2.10.2 Light pipes 33
2.10.3 Field lens 34
2.11 Field of View for Staring Arrays 34
References 35

Chapter 3. Primary Aberrations

3.1 Introduction 37
3.2 Primary Aberrations 37
3.2.1 Spherical aberration 37
3.2.2 Coma 38
3.2.3 Astigmatism 39
3.2.4 Field curvature 40
3.2.5 Distortion 41
3.2.6 Axial chromatic aberration 41
3.2.7 Lateral chromatic aberration 42
3.3 Calculations of Primary Aberrations 42
3.3.1 Spherical aberration 44
3.3.2 Coma 46
3.3.3 Astigmatism 47
3.3.4 Field curvature 47
3.3.5 Astigmatism and field curvature combined 48
3.3.6 Axial chromatic aberration 49
3.3.7 Numerical example 50
3.4 General Aberration Correction Methods 53
3.5 Doublets 53
3.5.1 Two elements, same material 53
3.5.2 Two elements, different materials 55
3.5.3 The achromat 56
3.6 Two Thin Air-spaced Elements 57
3.6.1 The Petzval objective 57
3.6.2 Refractive beam expanders 59
3.6.3 Telephotos 62
3.7 Reflective Optics 62
3.7.1 The spherical mirror 63
3.7.2 The Mangin mirror 65
3.7.3 Classical two-mirror configurations 67
3.7.4 The two-sphere Cassegrain system 68
3.7.5 The two-sphere Gregory system 71
3.7.6 Schwarzschild, a very special case 72
3.7.7 Reflective beam expanders 73
3.8 Diffraction Limit 74
3.9 Resolution of Imaging Systems 75
References 76

Chapter 4. Wave Aberrations

4.1 Introduction 77
4.2 Diverging and Converging Waves 77
4.3 Optical Path Length OPL 78
4.4 Optical Path Difference OPD (Wave Front Aberration) 78
4.5 Spherical Aberration 78
4.5.1 Numerical example 79
4.5.2 Best focus position 81
4.6 Third-Order Spherical Aberration 82
4.7 Depth of Focus 82
References 82

Chapter 5. Special Optical Surfaces and Components

5.1 Introduction 83
5.2 The Plane-Parallel Plate 83
5.2.1 Displacements 83
5.2.2 Optical micrometer 85
5.2.3 Aberration contributions 86
5.2.4 Application remarks 87
5.2.5 The wedge (thin prism) 90
5.3 Domes 91
5.4 The Ball Lens 93
5.4.1 Spherical aberration 94
5.4.2 An aspherized ball lens 94
5.5 Gradient Index Lens 95
5.6 Conic Sections and General Aspheres 97
5.6.1 Mathematical expressions 98
5.6.2 Reflectors with conic section surfaces 99
5.6.3 Lenses with conic section surfaces 100
5.6.4 Common two-mirror configurations using conic section surfaces 103
5.6.5 General aspheres (surfaces of rotation) 103
5.6.6 Two conic section mirrors with an aspheric corrector 104
5.6.7 Three-mirror configurations 105
5.7 Diffractive (Binary) Optics 107
5.7.1 The simple diffractive singlet 107
5.7.2 The hybrid achromat 109
5.7.3 Numerical examples 111
5.7.4 Diffraction efficiency 114
5.7.5 "Useful" spectral bandwidth 115
5.7.6 Diffraction efficiency for a particular order 115
5.7.7 The hybrid achromat, corrected for chromatic and spherical aberrations 116
5.7.8 Binary optics 117
References 119

Chapter 6. Design Examples

6.1 Introduction 121
6.2 Basic Assumptions for the High-and Low Temperature Applications 121
6.2.1 Optics for high-temperature system (3–5 μm) 122
6.2.2 Optics for low-temperature system #1 (8–12 μm) 122
6.2.3 Optics for low-temperature system #2 (8–12 μm) 123
6.2.4 Optics for low-temperature system #3 (8–12 μm) 123
6.3 The Improved Petzval Objective 125
6.3.1 Numerical example for a LWIR application 126
6.3.2 Manufacturing remarks 128
6.4 Instantaneous Field of View 128
References 129

Chapter 7. Thermal Effects

7.1 Introduction 131
7.2 Changing Parameters 131
7.3 Defocus with Change of Temperature 132
7.4 Defocus of Singlet 132
7.5 Athermalization with a Doublet 133
7.6 The Athermalized Achromat 135
7.6.1 The all-refractive athermal achromat 136
7.6.2 The hybrid athermal achromat 137
7.7 Cold Stop and Cold Shield 138
7.7.1 Cold stop 138
7.7.2 Cold shield 138
References 139

Chapter 8. Optical Coatings

8.1 Introduction 141
8.2 Effects at a Single Surface 141
8.3 Two Plane-Parallel Surfaces 142
8.4 Antireflection Coatings 143
8.5 Reflective Coatings 145
8.6 Typical Interference Filters 146
8.6.1 Angular sensitivity of filters 148
8.6.2 Thermal sensitivity of filters 149
References 150

Chapter 9. Image Evaluation

9.1 Introduction 151
9.2 Blur Spot Measurements 151
9.2.1 Circular mask 151
9.2.2 Slit 152
9.2.3 Knife Edge 152
9.3 Energy Distribution 154

9.4 Modulation Transfer Function 154
9.4.1 Overview 155
9.4.2 Contrast and resolving power 157
9.4.3 Diffraction MTF 160
9.4.4 Geometric MTF 161
9.4.5 Numerical example 161
References 163

Chapter 10. Diamond Turning

10.1 Introduction 165
10.2 Overview 165
10.3 Surface Finish 166
10.4 Scattering 168
10.5 Shape Correction 169
10.6 Optical Surface Testing 169
10.6.1 Surface roughness 169
10.6.2 Surface shape 171
10.7 Machining Time 171
10.8 Further Progress and Developments 172
References 172

Appendix

A.1 Paraxial Ray Tracing 173
A.1.1 Surface equations 173
A.1.2 Power equations (ray tracing through thin lenses) 175
A.2 Spherical Aberration of a Thin Lens 176
A.2.1 Derivation of expression 176
A.2.2 Blur spot size 178
A.3 Optical and Thermal Data for Some Infrared Materials 178
A.3.1 Selected materials for the 3–5-μm spectral band 179
A.3.2 Selected materials for the 8–12-μm spectral band 179
References 181

Index 181

PREFACE TO THE SECOND EDITION

The many positive remarks about the first edition and listening to the feedback from the students over the past five years encouraged me to expand upon the original material in this second edition.

To that end, the subject of beam expanders was modified, and achromats have been covered in more detail. In a chapter on Special Optical Surfaces and Components, the ball lens, gradient optics, and three-mirror configurations have been added.

New chapters are Wave Aberrations, Thermal Effects, Design Examples, and Diamond Turning. In Wave Aberrations, besides the concept, comparison of spherical aberration with the Seidel coefficient is discussed. The chapter Thermal Effects deals with methods of designing athermats, lenses that compensate for the undesired results caused by temperature excursions. The chapter Design Example is an application-based summary of the subjects covered in the earlier chapters of the book. Because diamond turning is especially suitable for producing aspheres and diffractive infrared elements, a chapter has been added to describe this manufacturing method.

To follow the style of the first edition, the added material contains practical approaches with approximations and many numerical examples.

I thank Rick Hermann and Sharon Streams of SPIE for their support and editorial assistance. I also thank my friend and colleague Robert E. Fischer for reviewing this second edition. His suggestions have been most valuable.

Max J. Riedl
February 2001

PREFACE TO THE FIRST EDITION

This tutorial is intended to provide a basic approach to the optical design of infrared systems. It is written for systems engineers whose expertise is outside the field of optics. The material presented can be applied directly to the initial optical system layout phase to evaluate trade-offs of various configurations. It will also be very helpful in conveying requirements and expectations to an experienced lens designer.

Over the past decades, much emphasis has been placed on the use of computers in lens design. Powerful programs developed for the lens designer have made it possible to explore new and different approaches for finding better solutions to optical design challenges. Unfortunately, the process of using computers to perform the required calculations is often referred to as *automatic lens design.* But to obtain a sensible optical system, one that is practical to manufacture and meets cost-related and other special demands, the starting configuration must have a chance to meet those demands.

Optimization programs only succeed in finding the best available solution from local conditions. Therefore, it is important to start in the right neighborhood. To identify this neighborhood—a promising starting point—is one of the goals of this tutorial text.

Designing lenses for the infrared region is in some ways easier than working in the visible spectrum, since the wavelengths are longer, the index of refraction of most lens materials is higher, and their relative dispersion is lower. This generally results in smaller primary aberrations. Third-order aberration calculations are often sufficient to predict meaningful performance expectations even if the system is simplified to a set of thin lenses. (lenses with zero thickness). The fact that the diffraction limit is 10 to 20 times larger in the infrared than in the visible region adds to the usefulness of applying third-order aberration theory.

Most of the materials used for infrared lenses and mirrors lend themselves to single-point diamond turning. For that reason, aspheric and diffractive surfaces are routinely employed since they are no more difficult to generate by this method than spherical surfaces and offer to the correction of aberrations. Aspheric and diffractive surfaces are discussed in detail and are also treated in several numerical examples.

Throughout this tutorial text, much emphasis has been placed on the practical aspect of the material presented. This is reflected by the many approximations that yield useful answers—especially welcomed during the proposal stage when time and resources are usually scarce.

This book is organized to follow the flow of radiant energy from the source to the detector. This flow is expressed with a simplified radiometric equation whose components identify the major contributors to overall system performance.

While a great deal of material is covered, details had to be limited. It is hoped that what is presented will be of value not only in the predesign stages of infrared systems, but also as a stimulus to dig deeper into the existing literature of this exciting field.

The material presented is a collection of notes from early in-house engineering seminars that led eventually to a more formally structured presentation as part of the SPIE short-course program. Much of what I am sharing in these pages is based on my long-time professional and personal association with Warren J. Smith, who along with Donald C. O'Shea reviewed the manuscript, and Lowell L. Baskins. My special thanks go to the reviewers for their substantial help; their comments and suggestions greatly improved the quality of the text.

Max J. Riedl
October 1995

HISTORICAL REMARKS

It was 200 years ago that Sir William Herschel, the Royal Astronomer to King George III, made his famous discovery. He wanted to find some protection for his eyes while looking into the sun. In his experiments he noticed an increased response as he scanned the thermometer from the blue end of the spectrum toward the red. This was not new; it had been done before. However, when he moved the thermometer into the dark portion beyond the red, where his eyes could not perceive any light, he noticed that the heat effect increased. That is the region we now call *the infrared region.*

Over the years, the thermometer was replaced by other detectors.[1] In 1829, the first thermopile was introduced by Nobili, which was improved four years later by Melloni. During the 1880s, several more sensitive detectors were developed. Most notable was the Langley bolometer, which was about 30 times more sensitive than Melloni's thermopile. In 1917, Case developed the first photoconductive detector, using thallous sulfide.

An interesting discovery occurred during the 1930s at the Institute of Physics at Berlin, when Kutzscher experimented with lead sulfide crystals he had found in Italy. Accidentally he noticed that these crystals were infrared sensitive.[2] This discovery led eventually to an infrared search and tracking device.

In 1952, the U.S Army built the first scanning thermal imagers, which were called *thermographs.*[3] With the development of cooled short-time-constant indium antimonide (InSb) and mercury doped germanium (Ge:Hg) photodetectors in the late 1950s, the first fast framing sensors appeared, and in 1956 the first long wavelength FLIR was built at the University of Chicago.

Now, the technology has advanced to a point where thermal imagers and other infrared devices have found their place in many applications other than military functions. These applications range from medical, border control, safety and security, telecommunication, forensic investigations, and many more.

References

1. R. Hudson, *Infrared System Engineering*, John Wiley (1969), p. 6.
2. A. R. Vogler, *The Odyssey of a Scientist*, California State University, Fullerton (1986).
3. J. M. Lloyd, *Thermal Imaging Systems*, Plenum Press (1975), p. 5.

CHAPTER 1

Radiometric Considerations

1.1 Introduction

This first part begins with identifying the basic elements that make up an optical system: source or target, aperture stop, field stop, image plane, entrance and exit pupils, entrance and exit windows. These elements and necessary radiometric definitions are discussed and applied in analyzing the extended simplified radiometric performance equation [Eq. (1.7)]. This equation is structured to include the major contributions to the detected signal in an optical system: the thermal radiation laws that describe the transmission through the atmosphere, the optics of the system, and finally the response of the detector.

These concepts are discussed in sequence, except the third one, which will be addressed last and in more detail, because it embraces the major subject of this tutorial, optics.

1.2 Basic Optical Relations

An axial ray travels from the axial object point through the lens and on to the image plane. The axial ray that passes through the edge of the aperture stop is called the marginal axial ray. The principal ray, also referred to as the chief ray, is an oblique ray from an off-axis object point through the center of the aperture stop. The marginal principal ray begins at the edge of the covered object and travels through the center of the aperture stop and the edge of the field stop. As will be discussed in Chapter3, these two marginal rays are the only ones needed to calculate the primary, third-order aberrations. u and u_p are the angles formed by the axial ray and the principal ray relative to the optical axis. This is indicated in Fig. 1.1. The aperture stop is the physical opening in the optical system that limits the size of the axial energy cone from the object. The image of the aperture stop in the object space is the entrance pupil, and the image of the aperture stop in the image space is the exit pupil. In lens systems, the object space is to the left of the first lens surface. The image space is to the right of the last lens surface. The opening that limits angle u_p of the principal ray is the field stop. Its image in object space is the entrance window. In the image space, the field stop image is the exit window. In Fig. 1.1, D is the size of the aperture stop and d' is the size of the field stop.

Figure 1.1 also indicates that an object located a distance s to the left of the lens is imaged at distance s' to the right of the lens.

In many cases, the object (target) can be considered to be located at infinity. The image is then formed in the focal plane of the lens. The axial ray angle u will be zero and the image distance s' becomes f, the focal length. The ratio of the focal length and the entrance pupil diameter is called the relative aperture or f-number ($f/\#$).

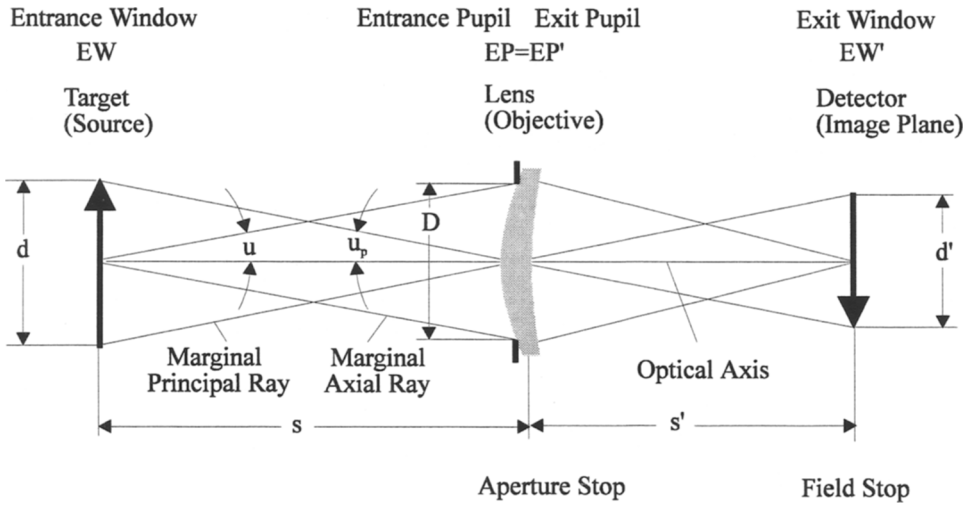


FIG. 1.1 Basic relations of imaging optics.

1.3 Signal-to-Noise Ratio

There are several means of evaluating the performance of a complete system, such as MRT[1] (minimum resolvable temperature) and MDT (minimum detectable temperature). However, because the objective of this tutorial is to concentrate primarily on the optical portion of an infrared system, we limit ourselves to just one measure, the signal-to-noise ratio (S/N).

In its simplest form, the signal-noise-ratio is stated by

$$S/N = \frac{P}{NEP}, \tag{1.1}$$

where P is the collected radiant power in watts that is received by the detector. *NEP* represents the noise equivalent power, a measure of the minimum signal that yields a unity signal-to-noise ratio.

The power can be calculated from

$$P = \frac{EP \times EW}{s^2} N = \frac{EP' \times EW'}{s'^2} N, \tag{1.2}$$

where *EP* and *EW* are the entrance pupil and entrance window areas (cm^2) and s is the separation of the entrance window from the entrance pupil (cm). N is called the radiance

of the source (W cm^{-2} $ster^{-1}$). The primed symbols refer to the image side of the system. EP' and EW' are exit pupil and exit window and s' is the separation of the two.

In this fundamental expression, N appears equal in both relations, indicating that no reduction of radiation has been accounted for due to loss in transmission or other factors.

The principal point to be made is that Eq. (1.2) is an invariant. It provides a choice for determining the power transfer from either the object side (target side) or the image side (detector side).

When the object is located at infinity, the image is formed in the focal plane. In this case, the area of the exit pupil is $D^2\pi/4$, and s' becomes f, which modifies the image-side expression of Eq. (1.2) to read

$$P = \frac{D^2 \pi d'^2}{4f^2} N \quad , \tag{1.3}$$

where d' is the linear dimension of the square detector. The detector is the exit window.

The radiated power in watts per square centimeter from a flat diffuse source surface into a hemisphere is the radiant emittance W. The relationship between the radiant emittance and the radiance is $N = W/\pi$.

With this and the substitution of $f/\#$ for f/D, Eq. (1.3) becomes

$$P = \frac{d'^2}{4(f/\#)^2} W. \tag{1.4}$$

The noise equivalent power NEP is a function of the detector size d', the electrical bandwidth Δf used in the measurement, and the detector figure of merit D^*, which has the somewhat unusual dimension of cm $Hz^{1/2}$ W^{-1}. D^* is the signal-to-noise ratio when 1 W is incident on a detector having a sensitive area of 1 cm^2, and the noise is measured with an electrical bandwidth of 1 Hz. So,

$$NEP = \frac{d'\sqrt{\Delta f}}{D^*}. \tag{1.5}$$

Substituting Eqs. (1.4) and (1.5) into Eq. (1.1) yields

$$S/N = \frac{D^* d' W}{4(f/\#)^2 \sqrt{\Delta f}} . \tag{1.6}$$

This simple expression indicates the strong influence of the chosen optical system. The S/N is inversely proportional to the square of $f/\#$, the relative aperture. This means that an IR system with an $f/1$ objective performs four times better with regard to S/N than

an $f/2$. Unfortunately, as we will see in Chapter 3, the "faster" (low $f/\#$) a lens is, the larger the aberrations.

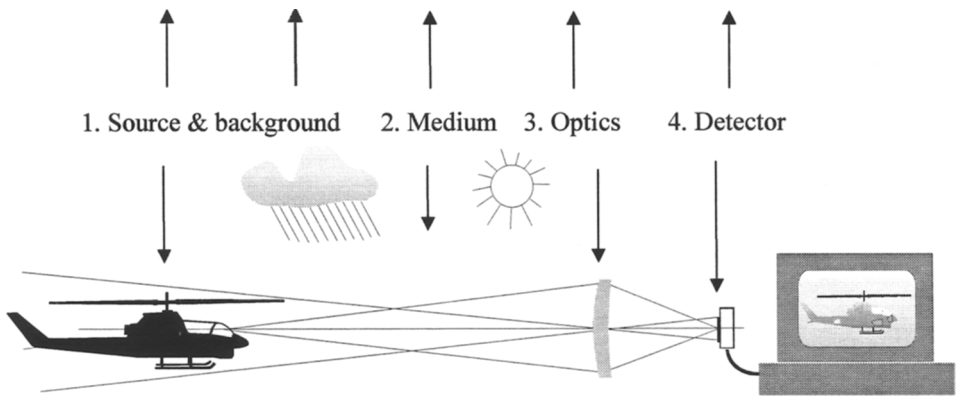
1.4 Extended Simplified Radiometric Performance Equation

Equation (1.6) can be extended to indicate the flow of the radiant energy from the source (target) to the detector by stating

Signal-to-noise ratio = [source power − background power]×[atmosphere transmission] ×[optical throughput]×[detector efficiency],

or

$$S/N = [W_T\varepsilon_T - W_B\varepsilon_B][\tau_A][\frac{\tau_O d'}{4(f/\#)^2}][\frac{D^*}{\sqrt{\Delta f}}]. \tag{1.7}$$



In Eq. (1.7), d' is included in the optics bracket. Even though d' is the linear dimension of the square detector, it is also the dimension of the field stop and exit window. The expression assumes that the image of the target is larger than the detector, i.e., the radiation received from the target "fills" the detector.

We shall now address each of the brackets of this simplified expression. Because of our emphasis on optics, the third bracket will be discussed last.

1.5 Thermal Radiation Laws

$$S/N = \boxed{[W_T\varepsilon_T - W_B\varepsilon_B]}\ [\tau_A][\frac{\tau_O d'}{4(f/\#)^2}][\frac{D^*}{\sqrt{\Delta f}}].$$

In this first bracket of Eq. (1.7), W is the radiant emittance in watts per square centimeter within the spectral bandwidth of interest. ε is called emissivity; ε is dimensionless and its meaning will be discussed shortly. The subscripts $_T$ and $_B$ refer to target and background respectively. Background radiation is the radiation coming from the surrounding target. The effective or net radiant emittance is the difference between

target and background radiation. This means that there has to be a radiation difference, positive or negative, to detect or resolve a target (object).

1.5.1 Blackbody radiation and Planck's law

A blackbody is defined as a perfect radiator which absorbs all radiation incident upon it.

In his investigation, to find a relation between the radiation emitted by a blackbody as a function of temperature and wavelength, Max Planck (1858–1947) developed the now famous equation named after him. His efforts laid the foundation of the quantum theory, for which he received the Nobel prize in 1918.[2]

On October 19, 1900, he first reported his findings, which were based on his experimental work.[2] Only two months later, on December 14, he presented the theoretical derivation of the equation[3] that described the blackbody radiation curve:

$$W_\lambda = \frac{C_1}{\lambda^5 \left[e^{C_2/\lambda T} - 1 \right]}, \tag{1.8}$$

where W_λ = spectral radiant emittance (W cm^{-2} μm^{-1}), λ = wavelength (μm), T = blackbody temperature (K), C_1 = 37,418, and C_2 = 14,388, when the area is in square centimeters. e = base of the natural logarithm (2.718.....).

In Eq. (1.8) notice there is a strong dependence on the wavelength and that W_λ goes to zero when $\lambda = 0$ and ∞. Planck's curve for a 500 K blackbody is shown in Fig. 1.2.

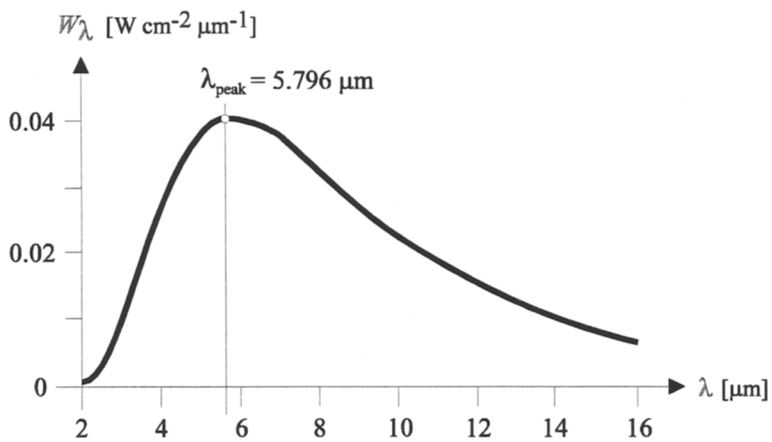


FIG. 1.2 Planck's curve for a blackbody source with T = 500 K.

To determine the radiant emittance over a spectral band, we integrate under the Planck curve between the limiting wavelengths λ_1 and λ_2 . An example of a selected spectral band is indicated in Fig. 1.3.

$$W = \int_{\lambda_1}^{\lambda_2} W_\lambda \, d\lambda \tag{1.9}$$

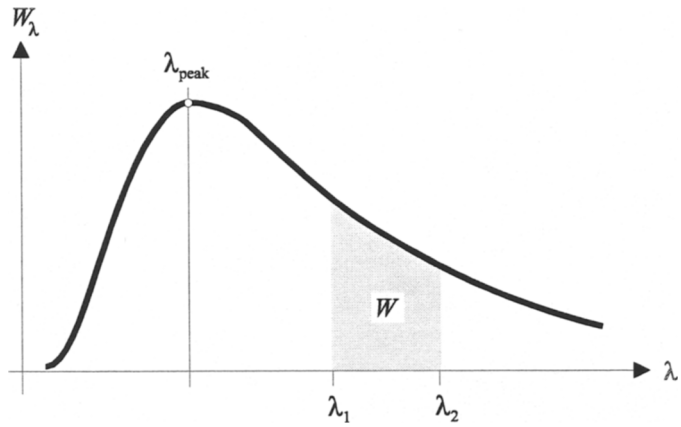


FIG. 1.3 Radiant emittance within a selected spectral bandwidth.

To perform the integration, one can use published tables or special slide rules[4]. A very convenient way is to apply Simpson's rule with the aid of a calculator. For that purpose we state that

$$W \cong \sum_{\lambda_1}^{\lambda_2} W_\lambda \, \Delta\lambda \ . \tag{1.10}$$

Using a $\Delta\lambda$ of 0.05 or even 0.1 μm is sufficient for summations over the mid-wavelength (MWIR) and long-wavelength (LWIR) infrared regions (3–5, and 8–12 μm). It is also good to remember that if the spectral band is relatively narrow, $W \approx W_\lambda \Delta\lambda$.

1.5.2 The Stefan–Boltzmann law

Independently, Josef Stefan (1835–1893) and Ludwig Boltzmann (1844–1906) found experimentally that the total radiation emitted by a blackbody (the area under the Planck curve) is

$$W_{\text{total}} = \sigma T^4, \tag{1.11}$$

where $\sigma = 5.66961 \times 10^{-12}$ (W cm^{-2} K^{-4}) and T = source temperature (K).

1.5.3 Wien's displacement law

There is a simple and interesting relationship between the peak wavelength and the temperature at which a blackbody radiates. Wilhelm Carl Werner Otto Fritz Wien (1864–1928), a Nobel prize recipient in 1911, discovered this behavior of the blackbody [Eq. (1.12)]. The equation states that the product of the peak wavelength and the source temperature is constant, which means the peak of the radiation shifts to shorter wavelengths as the temperature increases. Fig. 1.4 shows the displacement of the peak wavelength.

$$\lambda_p T = 2897.8 \quad (\mu\text{m K}). \tag{1.12}$$

Applying Planck's law, the spectral radiant emittance at the peak wavelength is

$$W_{\lambda\max} = 1.288 \times 10^{-15} T^5 \quad (\text{W cm}^{-2}\,\mu\text{m}^{-1}). \tag{1.13}$$

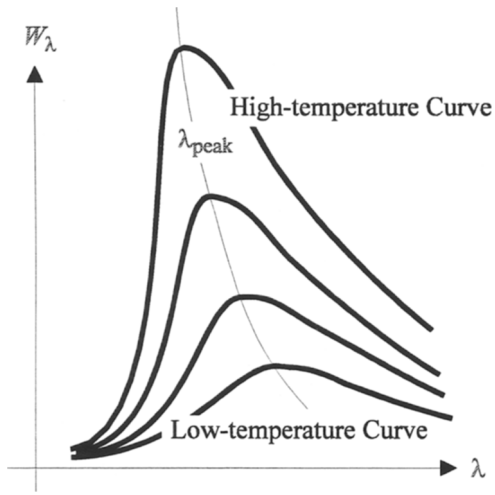


FIG. 1.4 Wien's displacement law.

1.5.4 Kirchhoff's law and emissivity

Gustav Robert Kirchhoff (1824–1887) stated in 1860 that "at thermal equilibrium, the power *radiated* by an object must be equal to the power *absorbed*." This leads to the observation that if an object absorbs 100 percent of the radiation incident upon it, it must reradiate 100 percent. As already stated, this is the definition of a blackbody radiator.

Most radiation sources are not blackbodies. Some of the energy incident upon them may be reflected or transmitted. The ratio of the radiant emittance W' of such a source and the radiant emittance W of a blackbody at the same temperature is called the emissivity ε of the source:

$$\varepsilon = W'/W . \tag{1.14}$$

With this relation, different types of radiation sources can be classified as indicated in Fig. 1.5 where the curve for the blackbody with $\varepsilon = 1$ is Plank's curve. The curve for a graybody is proportional to Plank's curve for all wavelengths. The spectral radiant emittance for a selective radiator varies not only with temperature but also with wavelength.

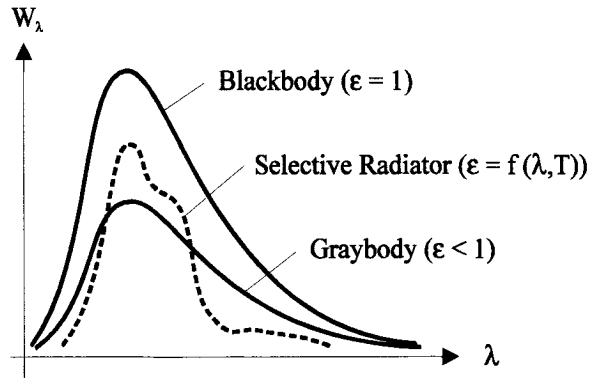


FIG. 1.5 Spectral radiant emittance of three types of radiators.

1.6 Transmission through the Atmosphere

$$S/N = [W_T\varepsilon_T - W_B\varepsilon_B]\boxed{[\tau_A]}\ [\frac{\tau_0 d'}{4(f/\#)^2}][\frac{D^*}{\sqrt{\Delta f}}] .$$

The second bracket of Eq. (1.7) addresses the transmission of radiant energy through the atmosphere.

To assess all influences that can affect the transmission is a very complex matter. Much work has been done by many researchers over many years to model the attenuation caused by the atmospheric gases.[5] Many variables, such as changes in temperature; pressure of the gases; shapes, sizes, and chemical compositions of suspended particles; and slanted optical path for example, make it extremely difficult to predict IR transmittance of the atmosphere. Our purpose here is to point out that when the IR system under consideration is to be employed over a long distance, atmospheric absorption effects may have to be included. Depending on the application, a cursory look may be sufficient. In some cases, however, a very detailed analysis will be necessary.

We limit our discussion to a horizontal path, near sea level, and develop approximations that allow us to obtain a basic understanding of the primary impact of the atmosphere on IR systems.

1.6.1 Permanent constituents of dry atmosphere

The relative concentrations of gases present in dry atmosphere are nearly constant and are therefore called the permanent constituents. Table 1.1 identifies these constituents and indicates their absorption characteristic in the infrared spectrum.

Table 1.1 Permanent constituents of dry atmosphere.[6]

Constituent (major)	Chemical formula	Percent by volume	Absorbs between 2 and 15 μm
Nitrogen	N_2	78.084	No
Oxygen	O_2	20.946	No
Argon	A	0.934	No
Carbon dioxide	**CO_2**	**0.032**	**Yes**
Neon	Ne	1.818×10^{-3}	No
Helium	He	5.24×10^{-4}	No
Methane	**CH_4**	**2.0×10^{-4}**	**Yes**
Krypton	Kr	1.14×10^{-4}	No
Nitrous oxide	**N_2O**	**5.0×10^{-5}**	**Yes**
Hydrogen	H_2	5.0×10^{-5}	No
Xenon	Xe	9.0×10^{-6}	No
Carbon monoxide	**CO**	**7.5×10^{-6}**	**Yes**

1.6.2 Variable constituents

The two major constituents in the atmosphere that vary with temperature and altitude are ozone and water vapor.

Since the maximum concentration of ozone occurs at a high altitude (between 10 and 30 km above the earth's surface), we can ignore its effect and state that

water vapor causes most of the absorption in the IR region at sea level.

1.6.3 Approximation (assumption)

Near sea level one can approximate the transmittance through the earth's atmosphere by considering the absorption effects of carbon dioxide and water vapor only. This is expressed by

$$\tau_{atmos} = \tau_{CO_2} \times \tau_{H_2O} \cdot \tag{1.15}$$

Here, H_2O represents water vapor whose density varies with temperature and humidity. Figure 1.6 shows the infrared wavelength regions where the transmittance is affected by the presence of water vapor and carbon dioxide.

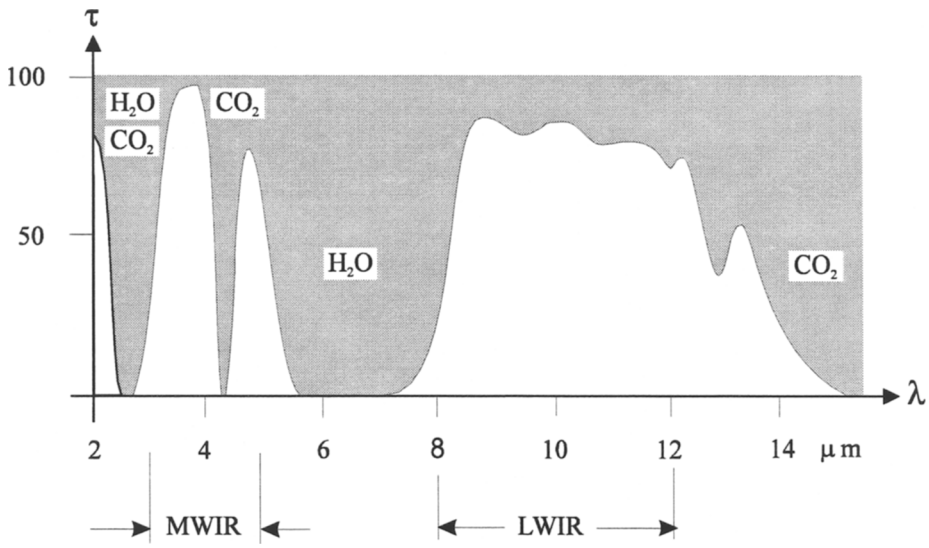


FIG. 1.6 Spectral windows in the infrared.

1.6.4 Precipitable water (definition)

The amount of water vapor contained in the optical path is called *precipitable water.* Important to remember is the word *vapor,* i.e., water in gaseous form. Precipitable water (Fig. 1.7) is defined as the depth of the layer of water that would be formed if all the water vapor along the line of sight was condensed in a container having the same cross section as the optical bundle.

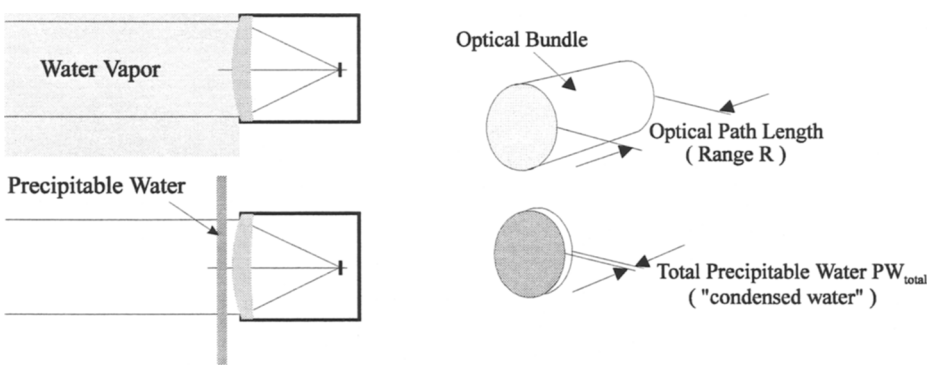


FIG. 1.7 Precipitable water, concept.

Notice that it is not necessary to know the cross-sectional area of the optical bundle (cylinder). If the cross section increases, more water is condensed, but it is spread over a larger area and its depth remains constant.

1.6.5 Humidity

The calculation of the total amount of precipitable water along the line of sight of an optical system includes the existing relative humidity. Let us state a few relationships about humidity in general and then apply them to an example. To simplify matters somewhat, we shall develop some approximations.

Absolute humidity [AH (gm^{-3})]

$AH(t)$ is the mass of water vapor in unit volume of the atmosphere at temperature t (°C). $AH_{sat}(t)$ is the maximum amount of water vapor the atmosphere can hold at temperature t. This condition is called saturation.

Relative humidity (RH)

Relative humidity is simply the ratio between the mass of water vapor per unit volume present in the air and the mass of water vapor in saturated air at the same temperature, i.e.,

$$RH = \frac{AH(t)}{AH_{sat}(t)} . \tag{1.16}$$

1.6.6 Precipitable water (calculation)

The amount of precipitable water is conveniently expressed in millimeter per meter path length and can be calculated by

$$pw = 10^{-3} AH(t) = 10^{-3} AH_{sat}(t) \times RH. \tag{1.17}$$

For the temperature range from 0° to 35°C, (32° to 95°F), an approximation for the absolute humidity at saturation is

$$AH_{sat} \cong 5.071962 e^{0.059688t} . \tag{1.18}$$

This approximation is accurate to better than 5% throughout the stated temperature range.

For a 1-km path length, Eq. (1.18) can be restated as

$$pw \cong 5.072 e^{0.0597t} \times RH \quad \text{(mm/km)}. \tag{1.19}$$

The nomogram shown in Fig. 1.8 is based on this expression.

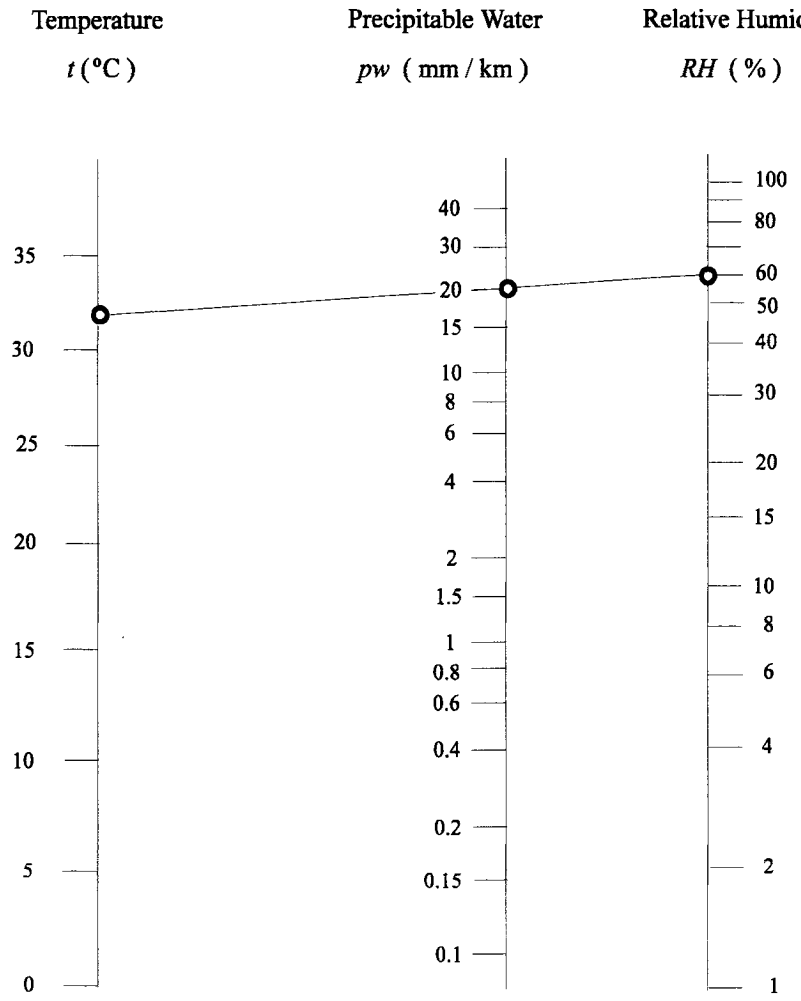


FIG. 1.8 Nomogram for determining the amount of precipitable water from the temperature and relative humidity.

In the example shown in Fig. 1.8, an atmosphere with 60% relative humidity at 32°C, contains 20 mm/km precipitable water.

1.6.7 Atmospheric transmission (calculation)

Equation (1.15) states that the total transmission through the atmosphere for specific conditions can be approximated by the product of the transmission components of carbon dioxide and water vapor.

To assess the effects over different spectral bandwidths, varying path lengths and precipitable water contents requires, even with our simplified expression, the aid of lookup tables, such as provided in Ref. 6. To determine the total transmission in the MWIR region (3–5 μm) for the example referred to in Fig. 1.8, we find the average

transmittance of CO_2 from the tables in Ref. 6 for a 1-km path length $\tau_{CO_2} = 0.848$, and the average transmittance for 20-mm precipitable water $\tau_{H_2O} = 0.591$. Therefore, the total transmission through the atmosphere over 1 km near sea level at 32 °C, and 60 percent relative humidity is approximately $\tau_{atmos} = 0.8.48 \times 0.591 \cong 0.50$. For the same temperature and humidity conditions, the total transmission over 1 km in the LWIR region (8 to 12 μm) is considerably better, namely $\tau_{atmos} = 0.995 \times 0.723 \cong 0.72$.

Figure 1.9 is a plot of the atmospheric transmission for both infrared windows (MWIR, and LWIR) over 1 km. The above calculated examples are indicated. The two curves provide information of the impact of the precipitable water contents variation over the two major IR windows.

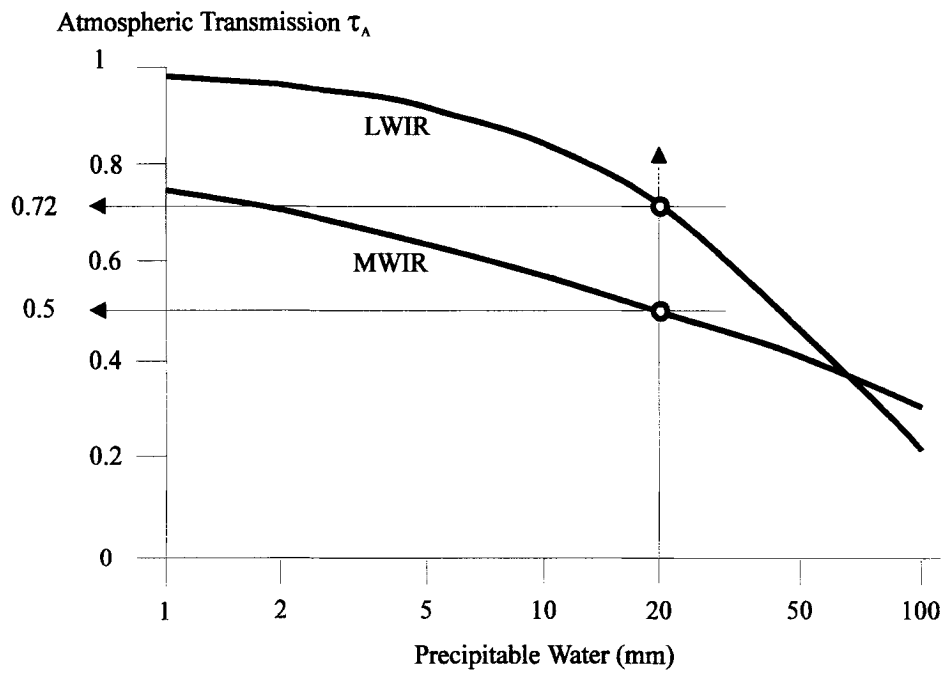


FIG. 1.9 Transmission over 1-km path for the MWIR and LWIR spectral bands.

1.6.8 Computer models

It is obvious that the subject of atmospheric transmission is extremely complex. The increase of available computer power and the dedicated efforts of many specialists in this field have led to computer models that are very detailed. At the present, some of the commercially available program models are *LOWTRAN*, *HITRAN*, and *SENTRAN*.

1.7 Typical IR Detectors

$$S/N = [W_T \varepsilon_T - W_B \varepsilon_B][\tau_A][\frac{\tau_O d'}{4(f/\#)^2}] \boxed{[\frac{D^*}{\sqrt{\Delta f}}]}$$

In this section we will briefly address the bracket of the simplified performance equation [Eq. (1.17)], which relates to the detector of the system. We begin by identifying the two basic groups of detectors that are commonly used for converting radiant energy into an electrical signal.

1.7.1 Thermal detectors

A thermal detector absorbs radiant energy, which causes a change of the detector's electrical characteristics. This electrical response to a change in the target temperature produces a signal that can be amplified and displayed.

One of the most attractive characteristics of thermal detectors is the equal response to all wavelengths. This contributes to the stability of a system that must operate over a wide temperature range. Another significant factor is that thermal detectors do not require cooling.

However, the response time of these detectors is in milliseconds and therefore relatively slow. Furthermore, their detectivity is as much as one or two orders of magnitude lower than that experienced with photon detectors.

The most common thermal detectors are thermocouple, thermopile, bolometer, and pyroelectric.

1.7.2 Photon or quantum detectors

Photon or quantum detectors operate on the quantum or photon effect. Photons are absorbed and produce free-charge carriers that change the electrical characteristic of the responsive element.

Photon detectors are much faster than thermal detectors; their response is in microseconds. As already mentioned, their detectivity is considerably higher. To obtain this high detectivity, the detector must be cooled. For moderate temperature reductions, one- or multistage thermoelectric coolers are employed. To provide cooling to very low temperatures of 77 K and even lower, cryogenic cooling methods must be applied.

1.7.3 Photoconductive detectors

This is the most widely used group of photon detectors. Their function is based on the photoconductive effect. Incident infrared photons are absorbed, producing free-charge carriers that change the electrical conductivity of the sensitive element.

Some of the materials found in this family of detectors are

• lead sulfide	PbS
• lead selenide	PbSe
• indium antimonide	InSb
• mercury cadmium telluride	HgCdTe .

1.7.4 Specific detectivity and noise equivalent bandwidth

One figure of merit used to describe the performance of a detector is the specific detectivity D^* (D-star). It is the signal-to-noise ratio when 1 W is incident on a detector having a sensitive area of 1 cm^2 and the noise is measured with an electrical bandwidth of 1 Hz. This definition was introduced by Clark Jones in the 1950s. Its dimension is a cumbersome cm $Hz^{1/2}$ W^{-1}; there has been the suggestion to change it to "Jones" units.[7]

The denominator of the bracket being discussed here refers the electrical bandwidth of the system's amplifier used for the noise measurement. The most common definition of this bandwidth is the frequency interval within which the power gain exceeds one-half of its maximum value. It is also often called the 3-dB bandwidth.[8]

Figure 1.10 shows the specific detectivity as a function of wavelength for some of the most commonly used IR photoconductive detectors.

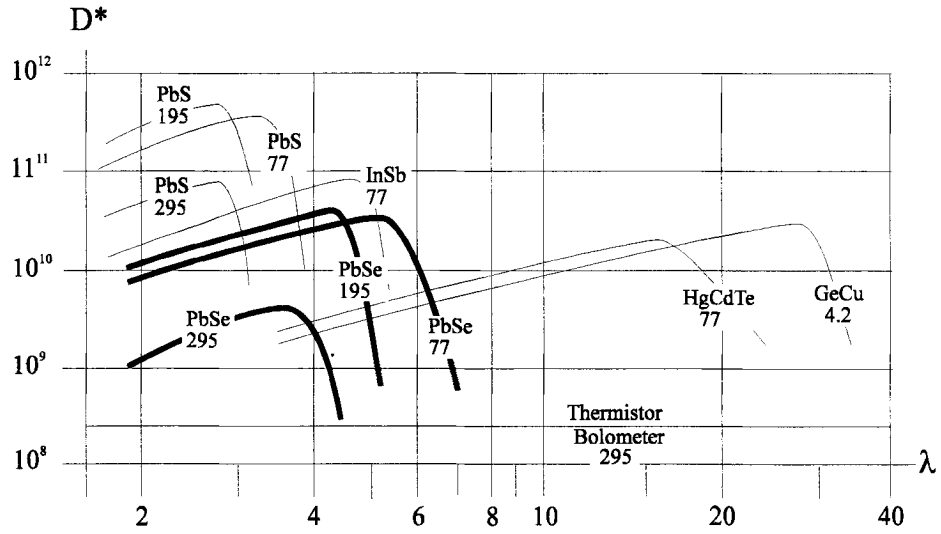


FIG. 1.10 Spectral sensitivity of some commonly used IR photoconductive detectors.

To illustrate the effect of cooling on performance, some curves in Fig. 1.10 have been highlighted. They refer to a lead selenide detector at three different temperatures; 295, 195, and 77 K (or 22,−78, and −196°C). It is interesting to note that the spectral roll-off response shifts to longer wavelengths as the temperature of the detector decreases. This is a very important property of a sensor. As a practical example, it can be seen clearly that to measure carbon dioxide, the sensitivity and stability is much greater if the

detector is cooled to 195 K than it would be for an uncooled detector. The CO_2 absorption occurs at 4.25 μm, which is where the cooled detector peaks. The importance is not only in the increased D^*; the detector is not operating at the slope of the curve, providing greater stability with temperature changes. It can also be seen that when the PbSe detector is cooled to an even lower temperature, the shift to longer wavelengths continues but there is no further gain in sensitivity (i.e., increase of D^*).

1.7.5 Detector configurations

Quite frequently, detectors are not single elements but are arranged in arrays. In thermal imaging, for example, a number of detector elements are mounted in either a single line or a two-dimensional layout. Linear arrays reduce the required scanning mechanism to one axis, horizontal or vertical. (Scanning is the process of scene dissection by sequentially detecting the radiation from a field covered by a single element or a linear array). Two-dimensional arrays eliminate scanning completely. This arrangement is known as staring arrays or focal plane arrays (FPA). One can easily imagine how advantageous such a grouping of many small detector elements in the focal plane can be.

From an optical standpoint, however, there is a trade-off, since such an arrangement presents an extended field of view which leads to a higher degree of complexity with regard to aberration corrections. This will be discussed in detail in Chapters 3 and 4.

Figure 1.11 shows a multistage thermoelectrically cooled detector array. The individual elements (pixels) are as small as 30 × 30 μm and the total number of elements in such an array can be more than 65,000.

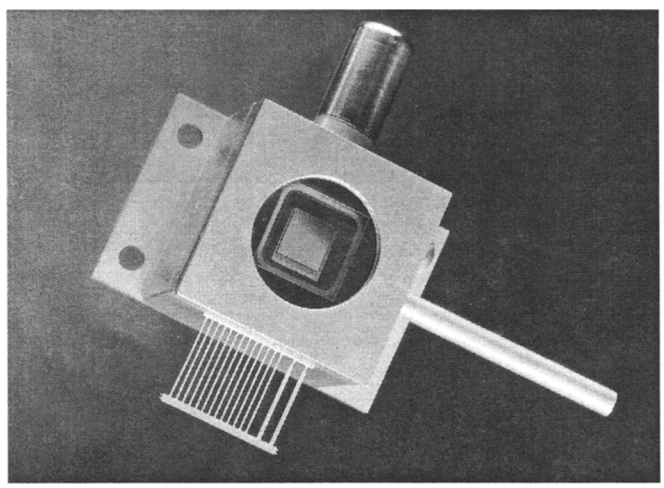


FIG. 1.11 Thermoelectrically cooled HgCdTe focal plane array. (courtesy of Hughes, Santa Barbara Research Center).

With the next chapter we shall begin to discuss the optics bracket of the extended simplified radiometric performance equation [Eq. (1.7)].

References

1. G. Holst, *Testing and Evaluation of Infrared Imaging Systems*, JCD Publishing Company (1993).
2. E. Hecht, *Optics*, second ed., Addison-Wesley (1990).
3. F. von Krbek, *Erlebte Physik*, Deutscher Verlag (1942), page 119.
4. *Infrared Radiation Calculator*, Infrared Information Analysis Center, Ann Arbor, MI, and EG&G Judson, Montgomeryville, PA and others.
5. F. Smith, *The Infrared and Electro-Optical Systems Handbook*, Vol. 2, *Atmospheric Propagation of Radiation*, ERIM & SPIE (1993).
6. R. Hudson, *Infrared System Engineering*, John Wiley (1969), page 118.
7. W. Rogatto, *The Infrared and Electro-Optical Systems Handbook*, Vol. 3, *Electro-Optical Components*, ERIM & SPIE (1993), page 251.
8. R. Hudson, *Infrared System Engineering*, John Wiley (1969), page 312.

CHAPTER 2

Basic Optics

2.1 Introduction

With this chapter, we begin to cover the third bracket of our simplified radiometric performance equation [Eq. (1.7)].

$$S/N = [W_T \varepsilon_T - W_B \varepsilon_B][\tau_A] \quad \boxed{[\frac{\tau_0 d'}{4(f/\#)^2}]}[\frac{D^*}{\sqrt{\Delta f}}],$$

where τ_0 is the net transmittance after absorption and Fresnel losses of all optical elements, which includes not only lenses and mirrors but windows and filters as well. As mentioned earlier, even though d' is the linear size of the detector element, it is included in the "optics bracket" because it is the dimension for the field stop.

Included in this chapter are some numerical examples of S/N calculations for the reader who is not concerned about optical aberrations and is satisfied with a very preliminary performance prediction for a conceptual system configuration.

But, at some point, more detailed analysis of the optics is required to do the job. For that purpose, a fundamental understanding of optics is essential.

We begin with the foundation of geometric optics, which is Snell's law, named after Willibrord Snel van Royen (1581–1626), a Dutch astronomer and mathematician who worked at the University of Leiden in Holland. It states that the sine-index product is equal across an interface when light passes from one transparent medium into another. The index of refraction is simply the ratio of the velocity of light in a vacuum to the light velocity in a medium. With reference to Fig. 2.1 we can write

$$\frac{\sin i}{\sin i'} = \frac{N'}{N}. \tag{2.1}$$

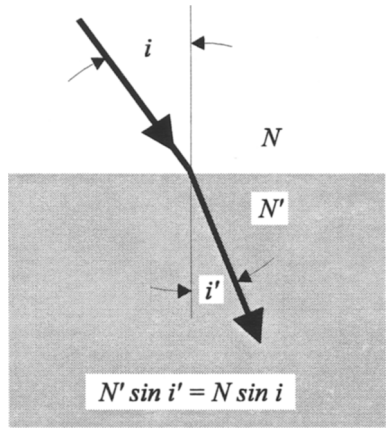


FIG. 2.1 Snell's law applied to a ray entering a denser medium ($N' > N$).

2.2 Snell's Law and the Prism

In applying Snell's law to the surfaces of a prism, we notice that a ray passing through the prism is always bent toward the thicker part, the base of the prism. This is illustrated by Fig. 2.2.

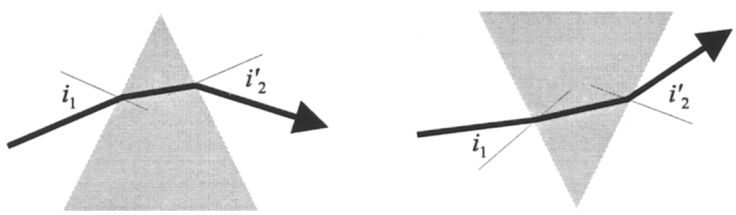


FIG. 2.2 Snell's law applied to a prism.

2.3 The Transition from a Prism to a Lens

If a number of truncated prisms are stacked as indicated in Fig. 2.3, rays emanating from an axial point converge or diverge symmetrically after passing through the prism array, depending on the orientation of the prisms. By increasing the number of prisms to infinity, a convex or concave cylindrical lens will emerge. For a rotationally symmetrical lens, the sections shown in Fig. 2.3 are meridional cross sections through the lenses.

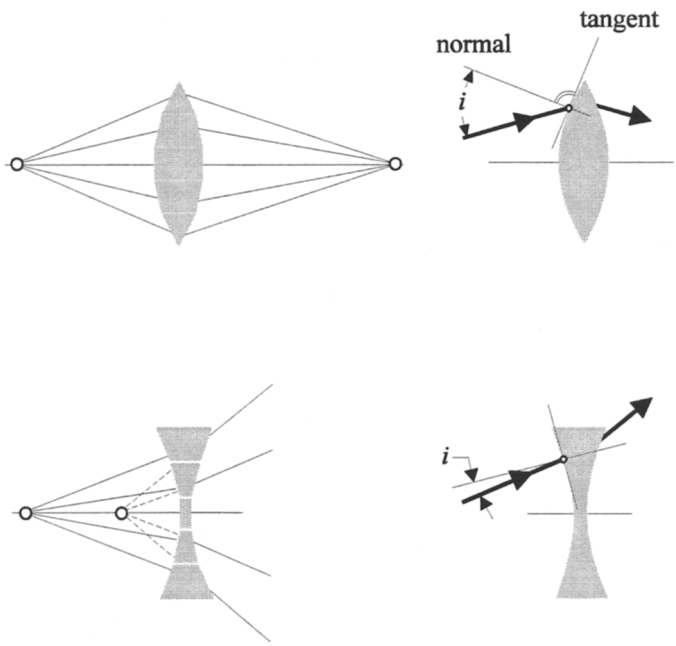


FIG. 2.3 Transition from prisms to lenses.

A tangent drawn where the ray enters the lens as shown in Fig. 2.3 is equivalent to the face of a "localized prism". Snell's law can be applied to the directional change of the ray, where i is formed by the ray with the normal to the lens surface.

2.4 Image Formation

For simplicity we limit this discussion to a single thin lens and use so-called first-order expressions.

A thin lens is one whose thickness is usually small compared to its focal length; in other words, it is assumed that the thickness is zero. We will show later that even though such a lens does not exist, the concept of a thin lens is very convenient. It simplifies calculations and still provides meaningful results.

First-order expressions are derived from the assumption that the sine of an angle is equal to the tangent and therefore equal to the angle itself. The domain for which this is the case is called the paraxial region, and the assumption is called the paraxial approximation because some simple approximations can be made in this region. This concept will be discussed in detail in Chapter 3.

It is important to pay close attention to the sign convention in optical ray tracing. Recently, there has been a change in the sign of the angle between a ray and the optical axis. Therefore, it is advisable to take a good look at the definition of the sign conventions used in a particular reference. The sign conventions shown in Fig. 2.4 are those most commonly used today.

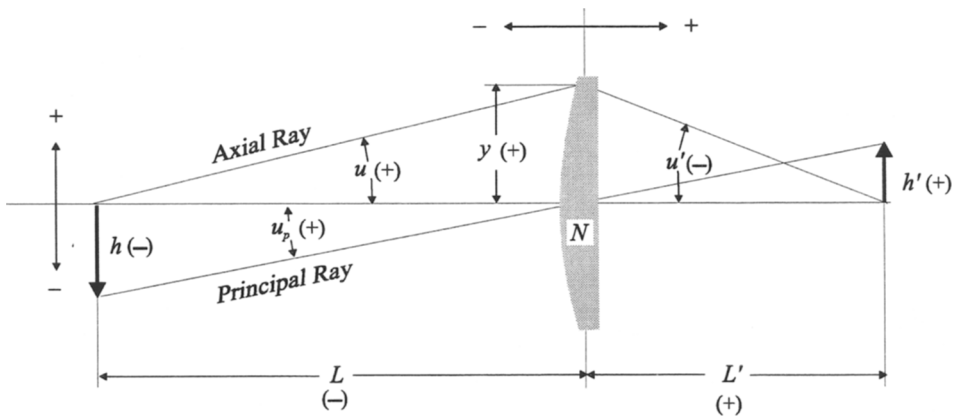


FIG. 2.4 Basic relations and sign conventions for a thin lens in the paraxial region.

If the object is located at infinity, angle u becomes zero and the distance from the lens where the image is formed is called the focal length f of the lens, and the image plane is the focal plane. The focal length of a converging lens is positive, and the focal length of a diverging lens is negative. This is indicated in Fig. 2.5. If the lens is surrounded by air, the focal lengths are equal for each side of the lens.

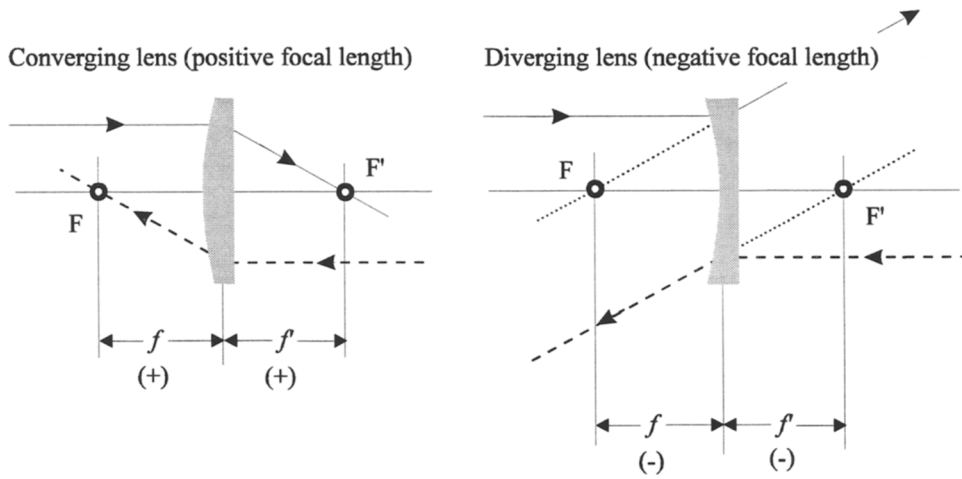


FIG. 2.5 Sign conventions for converging and diverging lenses.

There are a number of convenient and simple expressions to determine location, magnification, and orientation of an image for the arrangement shown in Figs. 2.4 and 2.5. The first expression is

$$\frac{1}{L'} = \frac{1}{L} + \frac{1}{f} . \tag{2.2}$$

Magnification can be stated in several ways:

$$m = \frac{h'}{h} = \frac{L'}{L} = \frac{u}{u'} . \tag{2.3}$$

The angular relation is often not recognized. The magnification is zero when the image is formed in the focal plane.

2.5 Object–Image Relations

The relationship between object and image is best understood by observing what happens to the image as an object, fixed in size, is being moved closer toward the lens (see Fig. 2.6). To find the location of the image and its size, a ray parallel to the optical axis is drawn from the top of object 1 (arrow head) to the right until the ray intercepts the plane of the lens. After passing through the thin lens, the ray travels downward and intercepts the optical axis at focal point *F'*. Another ray is drawn from the same object point aiming at focal point *F*. This ray leaves the lens parallel to the optical axis. Where the two rays intercept lies the arrow head's image point of image 1'. As the object—located to the left of the converging lens—is moved closer to the lens (object 2), the image 2' moves to the right, further away from the lens. When the object is placed at the focal point F of the lens, the image is formed at infinity. This is the principle of a collimator, whose function it is to simulate a source located at infinity. When the object is placed between focal point F and the lens, the formed image is virtual. 3' is the virtual image of object 3. To determine size and location of the virtual image, the exiting rays after passing through the lens are extended backward until they cross each other (dotted lines in Fig. 2.6). Object and image are at the same side of the lens and the image orientation does not change as in cases 1 and 2 where real images are formed. A real image can be received by a screen, a virtual image cannot. Good examples of virtual images are what can be seen in a mirror or through a magnifier.

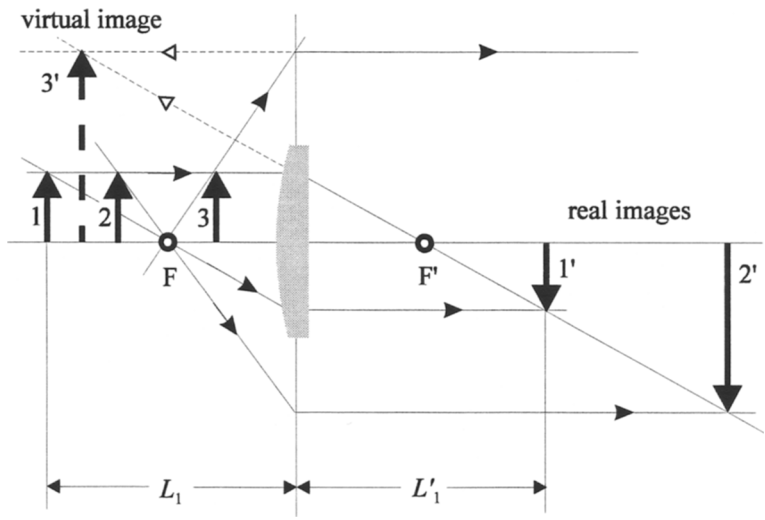


FIG. 2.6 Object–image relations.

It is important to understand the concept of real and virtual images, because as we will see shortly, entrance and exit pupils, which control the throughput of radiant energy through an optical system, can be real or virtual.

Applying Eqs. (2.2) and (2.3) to a lens with a focal length of $f = 30$ and an object height of $h = 10$, the three cases shown in Fig. 2.6 are tabulated in Table 2.1.

Table 2.1 Change of image location, magnification, and orientation with change of object position

	L	L'	m	h'	Image
1	−75	50	−0.667	−6.667	real
2	−45	90	-2	−20	real
3	−20	−60	+3	+30	virtual

Figure 2.7 shows the object–image relation [Eq. (2.2)] for a positive lens, plotted in focal length units. The three examples from Table 2.1 are identified.

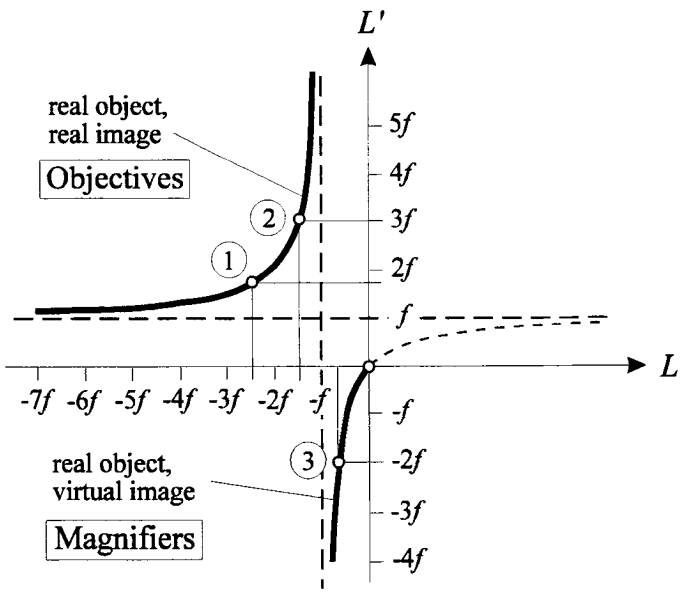


FIG. 2.7 Plot of Eq. (2.2) for a positive lens.

2.6 Stops, Pupils, and Windows

There are two stops that limit the bundles of energy passing through an optical system. One is the *aperture stop (AS)*, and the other is the *field stop (FS)*.

The aperture stop can be an element of its own, such as an iris, or it can be the edge of a lens mount. The aperture stop sets the angular limit of an axial ray. Its image in object space is called the *entrance pupil (EP)*, and its image in the image space is the *exit pupil (EP')*. Intermediate images of the aperture stop within the system are simply referred to as *pupils*.

For an optical system with a real object (target), the space extending to the left of the first optical surface is the *object space,* and the space to the right of the last optical surface is the *image space* if the image is real (detector plane). The situation becomes somewhat complex when the object or the image formed by the system is virtual.[1,2] These cases will not be considered here.

The opening that limits the size of an object that can be imaged by the system is called the *field stop*. The principal ray traveling from the edge of the field stop through the center of the aperture stop is the marginal principal ray. The field stop is the always related (conjugate) to the image and object planes. Frequently, the field stop lies in the image plane. In a camera it is the film, and in an infrared system it can be the detector itself. In a focal plane array (FPA), the full field is limited by the dimensions of the array. The field covered by an individual detector element is called the instantaneous field. The images of the field stop in object space and image space are respectively called *entrance window (EW)* and *exit window (EW')*. Intermediate images of the field stop are logically referred to as *windows*.

To better understand the interaction between the stops, we will discuss the arrangement shown in Fig. 2.8. The aperture stop is placed between two lenses and the detector is the field stop. From the layout, it can be seen that the size of the aperture stop *AS* controls the energy cone angle $2u$, and the field stop size *FS* limits the angular field size $2u_p$. The limiting rays for the energy cone and the field angle are the marginal axial and the marginal principal rays, respectively. Reducing the size of the aperture stop does not change the field coverage, nor does changing the field stop size have an impact on the cone angle $2u$.

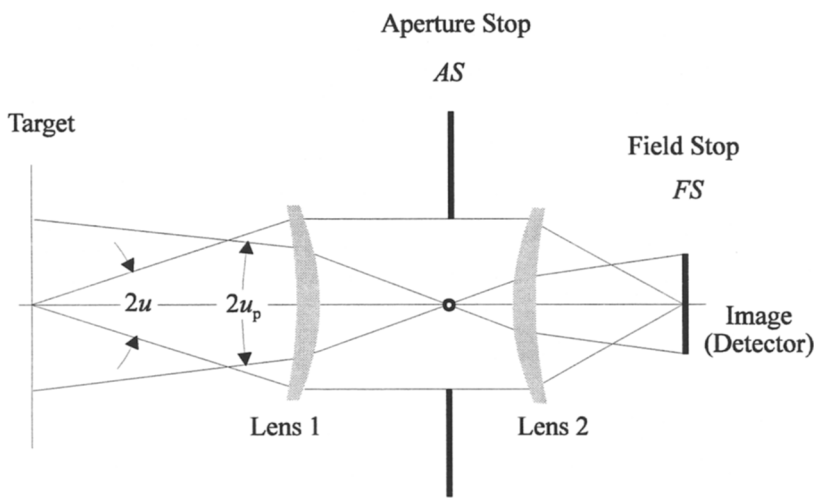


FIG. 2.8 Stops. The aperture stop controls $2u$; the field stop controls $2u_p$.

In applying the techniques of determining locations and sizes of images of the stops, the pupils and windows can be constructed. For the chosen example, both the entrance pupil and the exit pupil are virtual images of the aperture, as shown in Fig. 2.9. There is much more to be learned from Fig. 2.9.

From its start at the target, the marginal axial ray aims at the edge of the entrance pupil. The power of the first lens bends the ray to direct it toward the edge of the aperture stop. After passing through the second lens, the ray crosses the optical axis. This is the image location. Looking from the image side at the backward extended marginal axial ray reveals that this exiting ray appears to be coming from the edge of the exit pupil.

A similar observation can be made by tracing an oblique ray through the system from the edge of the object aimed at the center of the entrance pupil. This ray, the marginal principal ray, passes through the center of the aperture stop and terminates at the edge of the exit window. It appears to be coming from the center of the exit pupil.

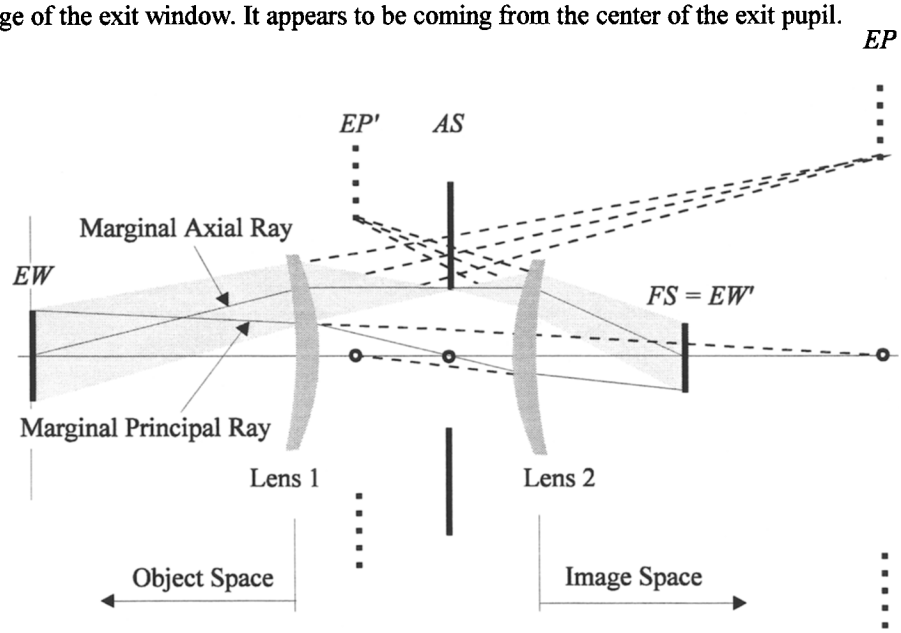


FIG. 2.9 Interaction among stops, pupils, and windows.

2.7 Throughput

Throughput is one name for the optical invariant that is used for the product of the pupil area and the solid angle subtended at this pupil by the window area. This means that the interaction between entrance pupil and the entrance window are the same as for the exit pupil and exit window. Stated mathematically,

$$\Gamma = \frac{EP \times EW}{S^2} = \frac{EP' \times EW'}{S'^2} = \text{Invariant}. \tag{2.4}$$

In this expression, *EP* and *EP'* as well as *EW* and *EW'* are the areas of the respective pupils and windows. *S* is the spacing between *EP* and *EW*, and *S'* is the separation of *EP'* from *EW'*. Figures 2.10 and 2.11 show the arrangement for our two-lens example.

Throughput has been called by many different names, optical extent, light gathering power, and area-solid-angle product. *Etendue* is the French term, and "geometrical light flux" is the translation of the German term *Geometrischer Lichtstrom*.[3,4]

What is really important is the choice that is available with this invariance. One can calculate the throughput using information from the object side or from the image side. When the image is at infinity, as in the case of a collimator, dimensions referring to the entrance pupil and entrance window can be used because *S* is a finite quantity, while *S'* is infinite. Similarly, when the object is at infinity, values for *EP'*, *EW'*, and *S'* from the image side can be used.

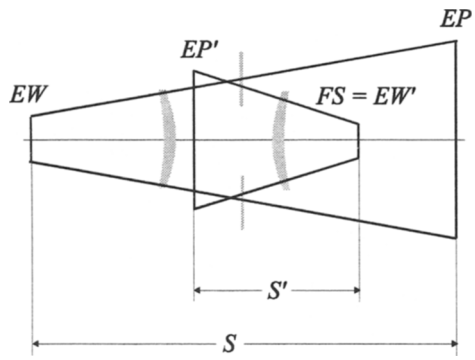


FIG. 2.10 Throughput invariance.

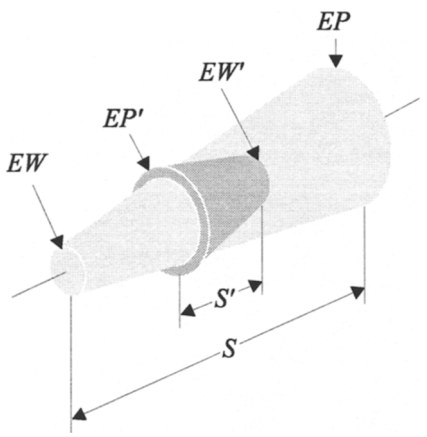


FIG. 2.11 A three-dimensional look at throughput invariance.

It must be added that pupils and windows are not always circular. In fact, for infrared systems, exit windows are mostly rectangular in shape. A typical example is the

staring focal plane array. In a Cassegrain telescope, the pupils are doughnut-shaped, due to the central obstruction of the secondary mirror. What counts is not the shape but the areas of the pupils and the windows.

2.8 Energy Transfer

The invariance Γ can be extended by multiplying it with the source radiance N to yield the power P received by the detector. As stated earlier, this is predicated on an overfilled detector that is the field stop.

$$P = \Gamma \times N = \frac{EP \times EW}{S^2} N = \frac{EP' \times EW'}{S'^2} N . \tag{2.5}$$

This is the second equation [Eq. (1.2)] that was presented in Chapter 1. To account for the transmission losses through the optics and, if applicable, for the atmosphere, the second equality for P has to be multiplied by the appropriate transmission factors. Those for the atmosphere were discussed in Chapter 1. Others will be explained in Chapter 8 on Optical Coatings.

2.8.1 Signal-to-noise calculations

We will now analyze a simple IR system. We begin with a high-temperature target and observe the impact on the signal-to-noise ratio (S/N) as the target temperature drops. We then switch to another wavelength region to see how that will influence S/N.

Let us assume a target temperature of T_T = 800 K and a background temperature of T_B = 300 K, and postulate that the emissivities for the target and the background are the same, namely $\varepsilon_T = \varepsilon_B = 0.85$. The chosen spectral band is from 3 to 5 μm, the linear lead selenide (PbSe) detector dimension d' = 0.05 cm (square), and its specific detectivity D^* = 0.5×10^9 cm $Hz^{1/2}$ W^{-1}. The total system transmittance $\tau = 0.6$, the lens diameter D = 5 cm, its focal length f = 20 cm, and the noise equivalent electrical band width Δf = 100 Hz.

With Eq. (1.10), we find for this high target temperature that the radiant emittance of the target W_T = 0.79105 W cm^{-2}. The background contribution W_B = 0.00059 W cm^{-2}. The relative aperture or $f/\# = f/D = 20/5 = 4$.

The S/N is therefore $S/N \approx \varepsilon \left(W_T - W_B\right) \dfrac{D^* d' \tau}{4\left(f/\#\right)^2 \sqrt{\Delta f}}$ = 15,747. This is a very comfortable number.

For the next calculation, the only change that will be made is to lower the target temperature from 800 to 320 K. This reduces the signal to noise ratio to 11.4, a very low number. Just as a point of reference, a signal-to-noise ratio of 100 would indicate that the noise is 1% of the signal.

The strong dependence between signal-to-noise ratio and target temperature is shown in Fig. 2.12. The curve is based on the above specified system with a 300 K background temperature.

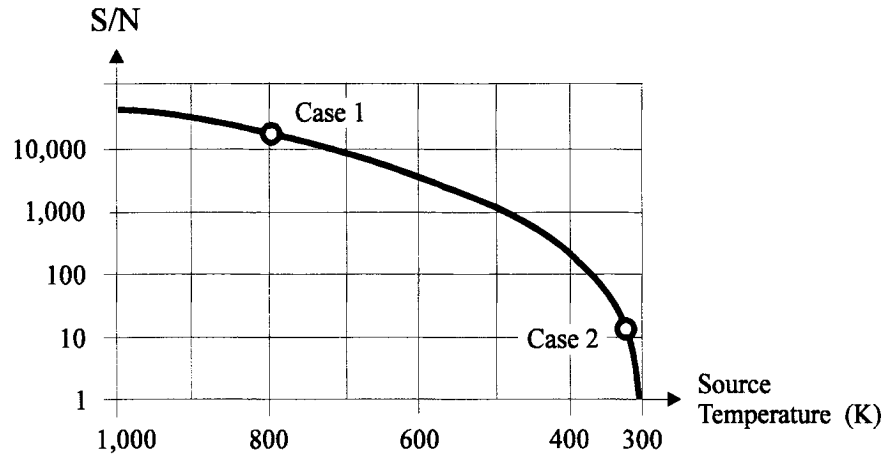


FIG. 2.12 S/N, a strong function of the source temperature.

What happens if we observe the target at longer wavelengths (8–12 μm)? Can we improve that low S/N of 11.4 for a target temperature of 320 K in a background with 300 K?

The radiant emittances in this spectral window for the given temperatures are $W_T = 0.0165$ and $W_B = 0.0121$ W cm^{-2}. These values result in a signal-to-noise ratio of 87.5. While this in an improvement by a factor of more than 7, there may be additional means for improvement. The most effective modification of the system would be to lower the f/#, because it enters into the equation inversely squared. In other words, a change from the assumed $f/4$ to $f/2$ would quadruple the S/N from 87.5 to 350.

This little exercise indicates how to read and interpret Eq. (1.7). It can save much time in evaluating a given situation during the initial layout phase.

2.9 Differential Changes

It is often interesting to know what differential changes in radiant emittance W can be expected with differential changes in temperature T. To refer to a temperature change of 1 K, we simply subtract the integrated Planck functions for two temperatures, 1° apart. Mathematically stated,

$$\frac{\delta W}{\delta T} = \int_{\lambda_1}^{\lambda_2} \frac{\delta W_\lambda(T)}{\delta T} dT \approx \sum_{\lambda_1}^{\lambda_2} W_\lambda(T+1)\Delta\lambda - \sum_{\lambda_1}^{\lambda_2} W_\lambda(T)\Delta\lambda \quad . \quad (2.6)$$

For $T = 300$ K: $\dfrac{\delta W}{\delta T_{3\text{–}5\mu}} = 2.1\times10^{-5}$ W cm^{-2} K^{-1} ,

and $$\frac{\delta W}{\delta T_{8-12\mu}} = 19.7\times10^{-5}\ \mathrm{Wcm^{-2}\,K^{-1}}.$$

Of course, as the temperature of the target changes, so does the ratio of the differential changes in the two spectral windows. We define this ratio as

$$\chi = \frac{(\delta W/\delta T)_{8-12\mu}}{(\delta W/\delta T)_{3-5\mu}}. \tag{2.7}$$

It can be clearly seen that there is a crossover point where the sensitivity becomes higher in the shorter wavelength region. That crossover occurs at 470 K, or about 387°F. This is illustrated in Fig. 2.13.

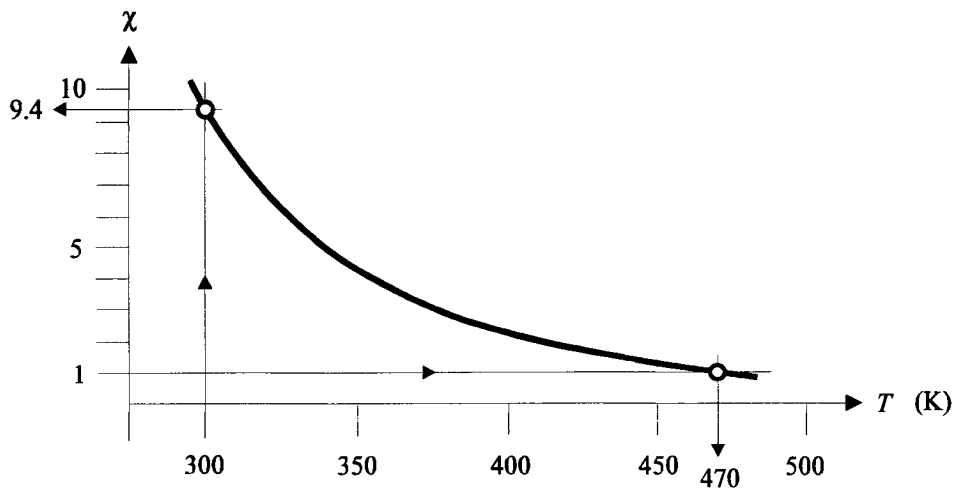


FIG. 2.13 Ratio of radiant emittances per degree Kelvin in the spectral regions 8–12 and 3–5 μm.

2.10 Optical Gain

Immersion lenses, light pipes, and field lenses can be employed to increase the S/N of an IR optical system without changing the field of view. These additional elements permit reduction of the physical size of the detector and have an added integrating effect over the sensitive detector area. As expected, there are certain limitations to such an arrangement.

2.10.1 Immersion lenses

As the name implies, the detector is immersed at the rear side of an optical element, i.e., the detector is in direct contact with the lens surface. If the immersion lens is a hemisphere, axial rays will not be refracted. Oblique rays, however, are bent as indicated in Fig. 2.14. The reduced detector size can be shown to be $\overline{d'} = d'/N$.

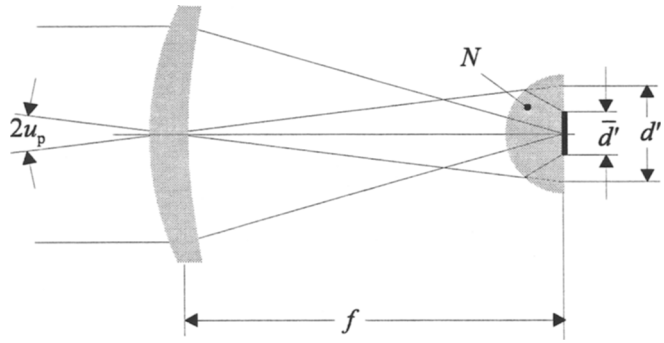


FIG. 2.14 Concentric hemisphere.

The radius of the hemisphere is not critical, because of the concentric arrangement; however, it should be of such a size that the angle between the extreme oblique ray and the normal at the lens is not too steep. The focal plane position is not affected by the insertion of the hemispherical immersion lens.

The signal-to-noise ratio is improved by a factor equal to the index of refraction N of the immersion lens. Therefore, the optical gain is N.

Another type of immersion lens is the hyperhemisphere. The optical gain with a properly shaped hyperhemispherical immersion lens is even higher, namely N^2. The proper shape refers to the so-called *aplanatic condition*, when aberrations are eliminated. For such a condition to exist, the following relations must be observed:

$$L = (N+1)R, \tag{2.8}$$

and

$$L' = \frac{(N+1)}{N}R. \tag{2.9}$$

With these relations, there will be a displacement of the focal plane as can be seen from Fig. 2.15.

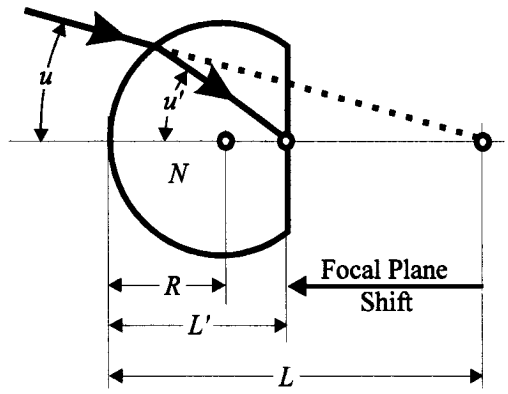


FIG. 2.15 Hyperhemispherical immersion lens.

The optical gain for this aplanatic case is equal to the square of the refractive index.

Theoretically, a gain of $4^2 = 16$ can be achieved with a germanium lens. However, there is a practical limit (see Fig. 2.16) which relates to the relative aperture or $f/\#$. For the aplanatic condition of the third kind as applied here, $\sin u' = N \sin u$. Therefore, at the limit, when $\sin u' = 1$,

$$\sin u_{max} = \frac{1}{N} = \frac{1}{\sqrt{4(f/\#)^2 - 1}} \tag{2.10}$$

or

$$(f/\#)_{min} = \frac{1}{2}\sqrt{N^2 + 1}\ . \tag{2.11}$$

A comparison of four materials that range from low to high refractive index shows that a careful approach is in order when a hyperhemispherical immersion lens is being considered to increase the S/N.

Material	N	$(f/\#)_{min}$ **(Limit)**
Glass	1.5	0.9
ZnSe	2.4	1.3
Si	3.4	1.8
Ge	4	2

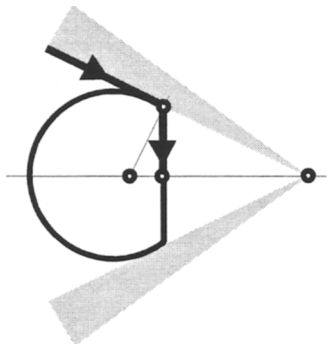


FIG. 2.16 A low $f/\#$ cone misses the lens.

Because of the steep angles involved at the extreme rays, it is advisable to reduce the limit even further. A factor of two may be reasonable.

2.10.2 Light pipes

The type of light pipes discussed here are solid and hollow cones that find their application in reducing the detector area. These cones are not imaging elements, they concentrate the received energy from the target onto a smaller cross section. In the process they have an additional effect; the energy is integrated, compensating for any existing non-uniform response across the detector area. Figure 2.17 shows the basic arrangement. Some possible variations of the light pipe are indicated in Fig. 2.18. Solid cones with a flat or curved front surface are also suggested. Further configurations are parabolic reflectors and a combination of refractive and reflective elements. A more detailed discussion of such light pipes or cone condensers can be found in Ref. 5.

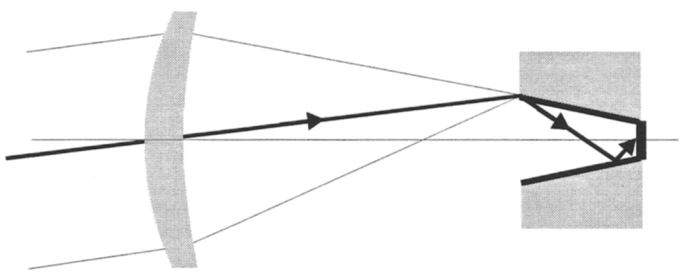


FIG. 2.17 Reflective light pipe (cone condenser).

It is apparent that the optical gain, which is the ratio between the entrance area of the light pipe and the detector area, is limited. If the cone angle of the pipe is too steep, the entering ray will reverse its direction and never reach the detector. In such a case, the light pipe acts as a retroreflector.[6]

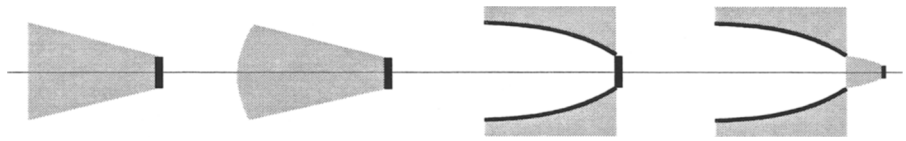


FIG. 2.18 Some possible shapes of light pipes to increase S/N.

2.10.3 Field lens

A lens placed at or near the focal plane of the objective lens can also serve to reduce the needed detector size for a given field coverage. The power of the lens is chosen to image the aperture stop of the objective into the detector plane. This arrangement has the same integrating effect as the cone already discussed. The lens used for the reduction of the detector is called the *field lens*. If placed exactly at the focus of the objective, it has no effect on the power of the objective. For a high optical gain factor, the field lens consists of several elements to correct for aberrations. From Fig. 2.19, it can be concluded that the optical gain is the ratio between the areas of the field lens and the detector.

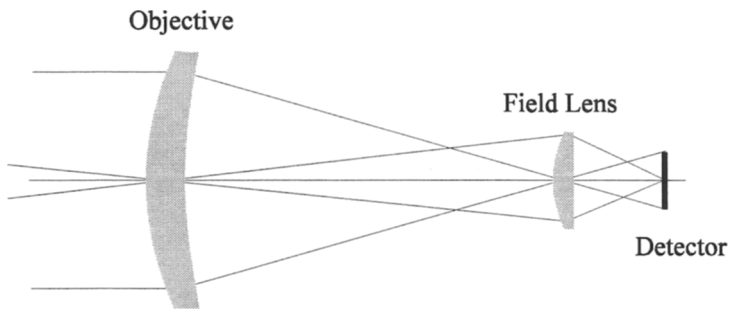


FIG. 2.19 Field lens images aperture stop onto detector .

2.11 Field of View for Staring Arrays

Frequently, the field of view for staring focal plane arrays is expressed in horizontal and vertical dimensions as indicated in Fig. 2.20. It must be remembered that from an optical standpoint, the total field dimension is the diagonal of the array. The extreme off-axis dimension is the one that enters into the aberration expressions. It is also the extent that needs to be clear of any obstruction.

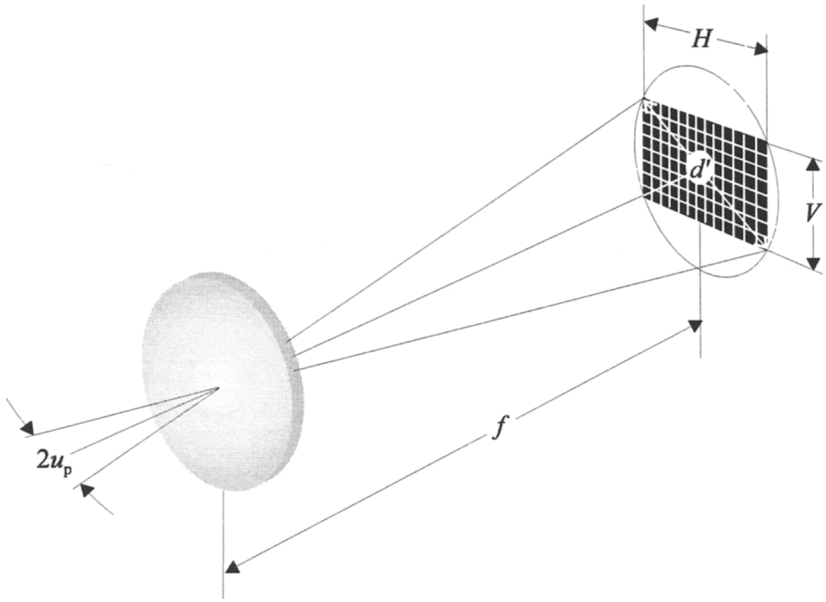


FIG. 2.20 Optical field of view $d' = \sqrt{H^2 + V^2}$.

References

1. R. Kingslake, *Optical System Design*, Academic Press (1983), page 10.
2. Bergmann, Schaefer, *Lehrbuch der Experimentalphysik,* Band III, F. Mantossi, *Optik,* vierte Auflage, Walter De Gruyter & Co (1966), page 78.
3. W. Steel, "Luminosity, throughput, or *etendue*?", Applied Optics 13, (1974), page 704.
4. K. Räntsch, *Die Optik in der Feinmesstechnik*, Carl Hanser Verlag (1949), page 53.
5. D. E. Williamsson, "Cone channel condenser optics", J. Opt. Soc. Am. 42, (1952), page 712.
6. W. J. Smith, *Modern Optical Engineering*, McGraw-Hill, Inc. (2000), page 280.

CHAPTER 3

Primary Aberrations

3.1 Introduction

In the previous chapters, comments were made relating to aberrations. Now it is time to discuss and explain them and find out how they affect the performance of an optical system.

In general, aberrations are deviations of an image point from its ideal position. These aberrations come in various forms. For this fundamental tutorial text we limit ourselves to the so-called primary aberrations:

- Spherical aberration
- Coma
- Astigmatism
- Field curvature
- Distortion
- Axial chromatic aberration
- Lateral chromatic aberration.

These seven aberrations will first be described and then expressions will be presented to be used for determining the quantitative impact of each on image quality.

3.2 Primary Aberrations

One can separate these aberrations by general categories. The first five types listed deal with monochromatic radiation; the last two address polychromatic effects. Sometimes the categories are split between on-axis and off-axis aberrations. By this definition, spherical and axial chromatic aberrations are on-axis aberrations because they refer to object points located on the optical axis (the system's axis of symmetry). The rest are off-axis aberrations.

3.2.1 Spherical aberration

Applying Snell's law across a spherical surface of a lens or a mirror shows that rays closer to the edge of such an element are bent more strongly than they should to meet with rays closer to the center of the element at the optical axis crossover point of these rays. However, because spherical surfaces are much easier to manufacture than aspheric ones, they are by far the more standard surface to be found in optical systems.

Figure 3.1 indicates that the marginal axial ray, after passing through the lens, intercepts the optical axis at point M. The ray offset from the optical axis by an infinitesimal amount (a ray in the paraxial region) intersects the axis at P. If the rays enter

the lens parallel to the axis, P is the focal point of the lens, the location of the paraxial image plane. Deviations from the paraxial image plane are the measure of spherical aberration. According to this statement, the distance PM in Fig. 3.1 is the *longitudinal spherical aberration.*

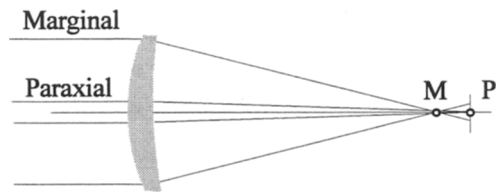


FIG. 3.1 Spherical aberration is the variation of focus with aperture.

3.2.2 Coma

Figure 3.2 shows two oblique rays, A and B, entering the edge of the lens as part of a bundle of light, parallel to the principal ray $\overline{PP}$. Rays A and B cross the image plane at a different height than the principal ray. The difference, $h'_m - h'_p$, is a measure of coma.

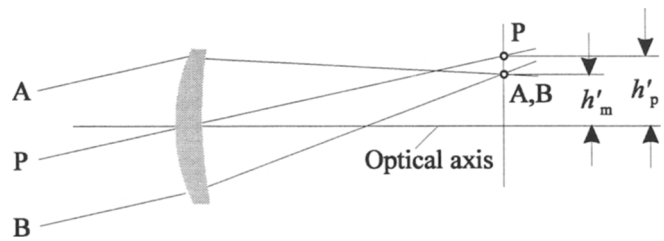


FIG. 3.2 Coma is the variation of magnification with aperture.

The image pattern formed by tracing many rays of the entering bundle through the lens is shown in Fig. 3.3. The optical axis and the principal ray of the system are identified by PO and PP, respectively. The construction of the image blur shape is presented by the sequence (a) through (e) The circle marked with point pairs 1–1, 2–2, 3–3, and 4–4 is the exit pupil of the lens. Rays entering the lens, parallel to the principal ray, passing through these identified points are focused in the image plane as points 1, 2, 3, and 4, forming the outline of the image blur that resembles the shape of a comet, hence the name *coma.* The plane that contains points 1 from the exit pupil and the principal ray PP is called the tangential plane. The plane perpendicular to the tangential plane, containing the principal ray and exit pupil points 2 is the sagittal plane. The image pattern [Fig. 3.3(f)] identifies the distance P–2 as sagittal coma and P–1 as tangential coma. As noted, the tangential coma is three times the size of the sagittal coma.

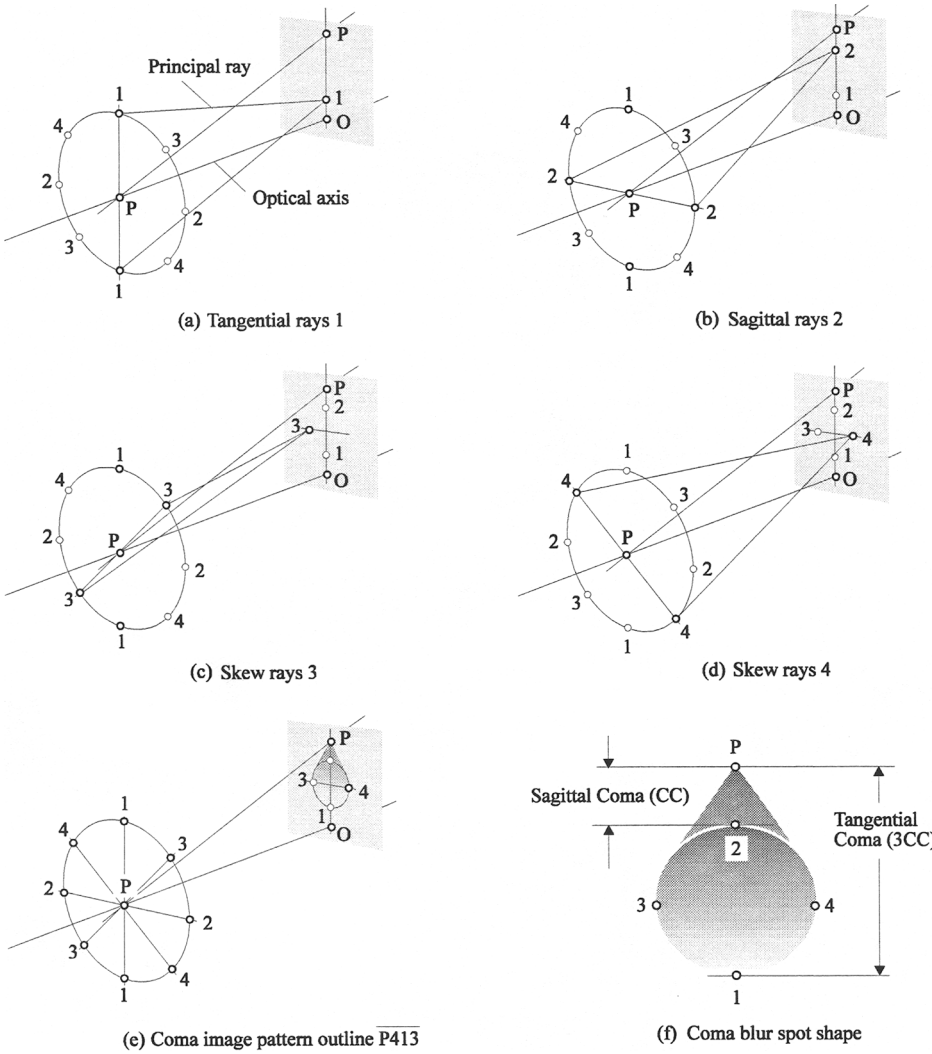


FIG. 3. 3 Construction of coma blur spot.

3.2.3 Astigmatism

Astigmatism is best described with a three-dimensional picture. Figure 3.4 identifies two planes that are perpendicular to each other. The plane that contains the object point and the optical axis is called the meridional or tangential plane. The rays in this plane form the tangential fan. The plane perpendicular to the tangential plane, containing the object point, is the sagittal plane.

On the image side there is a longitudinal separation between the tangential and sagittal images. The images of an object point are formed in two different planes. Because

of this, the sagittal and tangential images of the object point are lines. Approximately halfway between these two image lines lies the so-called circle of least confusion, a blur with the smallest linear dimension as its diameter.

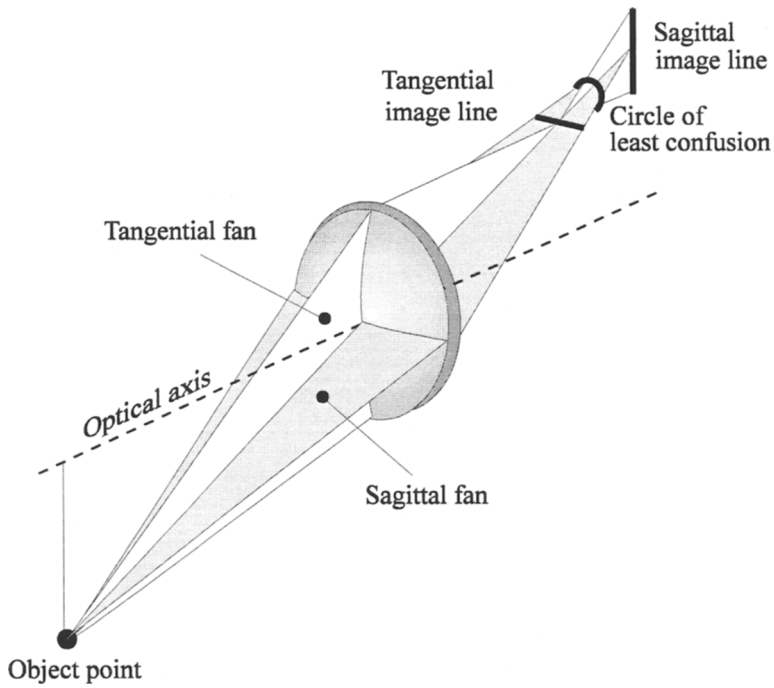


FIG. 3.4 Astigmatism.

3.2.4 Field curvature

Even if spherical aberration, coma, and astigmatism were absent, the image of an off-axis object point would still not lie in a plane without correction for field curvature. Assuming the first three aberrations were eliminated, the image surface would be parabolic in shape. Depending on the power of the lens, the parabola is either curved toward the lens or away from it. Positive lenses (positive-power lenses) have an inward-curved image surface (toward the lens). Negative lenses (negative-power lenses) have their image surface bent in the other direction. This basic field curvature is named after Josef Max Petzval, who discovered the behavior of astigmatism relative to field curvature. Petzval was a Hungarian mathematician who died heartbroken after burglars destroyed much of the manuscript in which he had recorded, among other findings, his derivations of aberration coefficients up to the seventh order.[1] The image surfaces and the Petzval surface are identified in Fig. 3.5. Measured parallel to the optical axis, for the single element shown, the tangential image surface is three times farther away from the Petzval surface than the sagittal image surface. If astigmatism is eliminated, the tangential and sagittal image surfaces coincide and lie on the Petzval surface.

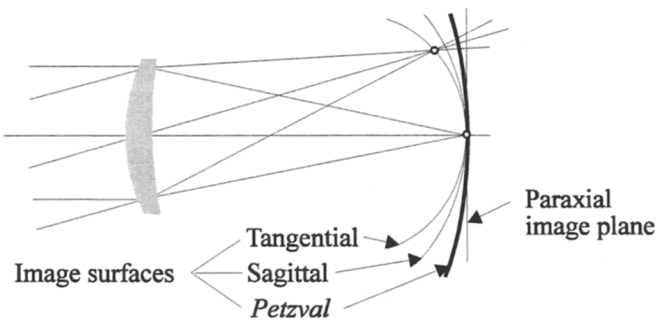


FIG. 3.5 Field curvature and Petzval surface.

3.2.5 Distortion

Distortion is the variation in the magnification of the optical system with field angle. If the magnification increases with the field, pincushion distortion results. If it decreases, barrel distortion is present. This is depicted in Fig. 3.6. Pincushion distortion is called positive; barrel distortion is negative.

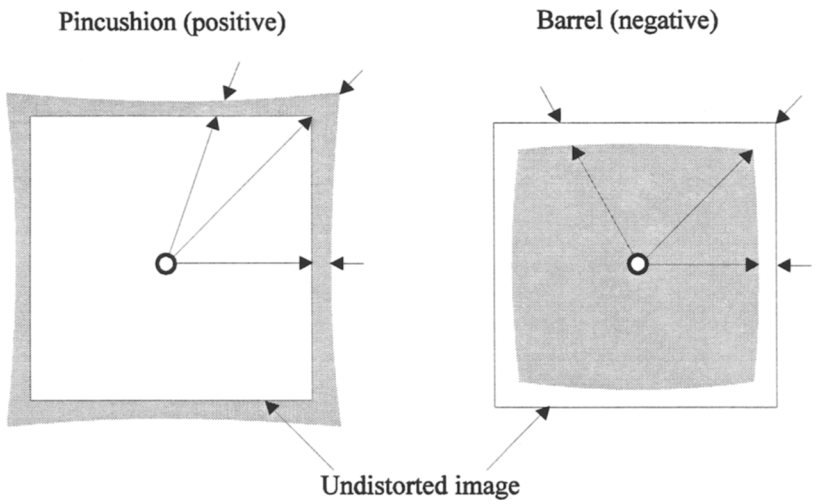


FIG. 3.6 Distortion.

3.2.6 Axial chromatic aberration

As stated in Chapter 2 under Snell's law, the amount of ray bending depends on the index of refraction. Because the index of refraction is not constant but varies with wavelength, a lens focuses rays of different colors at different places along the optical axis. This difference or spread between the two focus positions is called longitudinal axial chromatic aberration. For a single positive lens, the short-wavelength focus is closest to the lens (see Fig. 3.7).

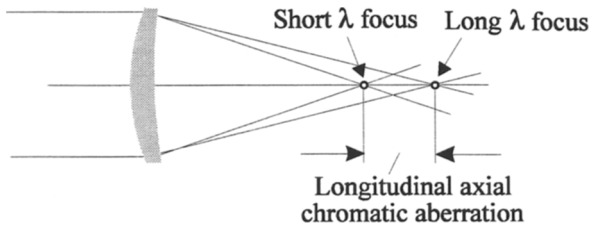


FIG. 3.7 Longitudinal axial chromatic aberration.

3.2.7 Lateral chromatic aberration

The separation of an imaged off-axis point, formed by rays of different colors, is called lateral chromatic aberration or lateral color. It is a consequence of the difference in magnification with wavelength. Figure 3.8 illustrates this situation.

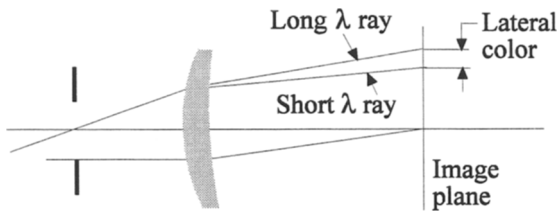


FIG. 3.8 Lateral chromatic aberration or lateral color.

3.3 Calculations of Primary Aberrations

To determine the aberrations described, we apply expressions developed with the so-called Seidel aberration theory and neglect the effects of the thickness of the lens.[2] Seidel's theory is also named third-order theory because it is based on the truncation of the series expansion for the sine of an angle after the second term, as indicated in Eq. (3.1).

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots \tag{3.1}$$

The power of a lens is the reciprocal of its focal length, expressed as

$$\phi = \frac{1}{f} = (N-1)\left[\frac{1}{R_1} - \frac{1}{R_2} + \frac{(N-1)t}{NR_1R_2}\right], \tag{3.2}$$

where R_1 and R_2 are the radii of the first and second surfaces respectively, t is the thickness of the lens, and N is the refractive index of the lens material as indicated in Fig. 3.9. The radius is positive if its center lies to the right of its surface; if it's to the left, it is negative. Therefore, R_1 in Fig. 3.9 is positive; R_2 is negative.

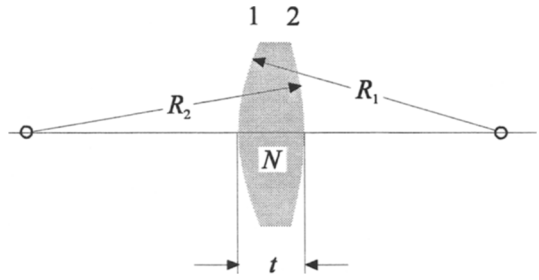


FIG. 3.9 The thick lens.

By neglecting the thickness t of the lens, Eq. (3.2) reduces to

$$\phi = (N-1)(c_1 - c_2) = (N-1)c, \tag{3.3}$$

where c_1 and c_2 are the curvatures or reciprocal radii of the lens surfaces. c is called the net or total curvature of the lens. Equation (3.3) states that as long as the net curvature is kept constant, the surface curvatures can vary without changing the power of the lens. This changing of the lens shape is called lens bending. It is a powerful basic tool in lens design, because the shape of the lens affects spherical aberration and coma but not its focal length. The dependence of spherical aberration on the shape for a germanium lens ($N = 4$) is illustrated in Fig. 3.10, where spherical aberration is plotted with respect to the ratio of the first surface curvature and the net curvature, c_1/c. This ratio is called the shape factor K.

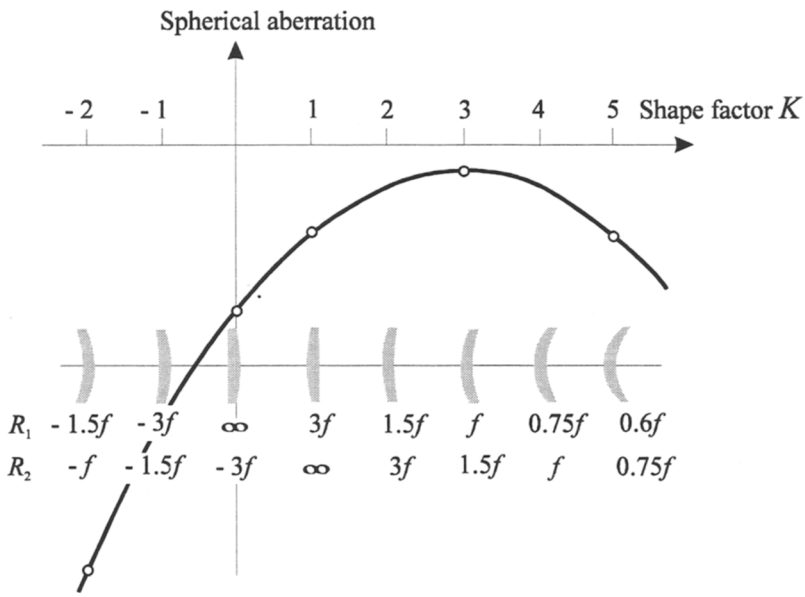


FIG. 3.10 Change of spherical aberration with change of lens shape (lens bending) for a germanium lens with index $N = 4$. Minimum spherical aberration is obtained in this case when the first radius of the lens is equal to its focal length and the second radius is 1.5 times the focal length ($K = 3$).

3.3.1 Spherical aberration

The angular blur spot size in radians is

$$\beta_{\text{spher}} = \frac{B}{f} = \frac{N^2 - (2N+1)K + \left[(N+2)/N\right]K^2}{32(N-1)^2(f/\#)^3}, \tag{3.4}$$

where N is the index of refraction of the lens material and K the shape factor discussed above. f is the focal length, and $f/\#$ is the relative aperture.

The best shape for a singlet is obtained by differentiating Eq. (3.4) and solving for K. In doing so, one finds

$$R_1 = \frac{2(N+2)(N-1)}{N(2N+1)} f, \tag{3.5}$$

and

$$R_2 = \frac{2(N+2)(N-1)}{N(2N-1)-4} f. \tag{3.6}$$

A lens shaped as suggested by these equations exhibits a minimum angular blur spot size of

$$\beta_{\text{spher}_{\text{min}}} = \frac{N(4N-1)}{128(N-1)^2(N+2)(f/\#)^3}. \tag{3.7}$$

The minimum linear blur spot B is located a distance Δ away from the paraxial image plane (see Fig. 3.11).

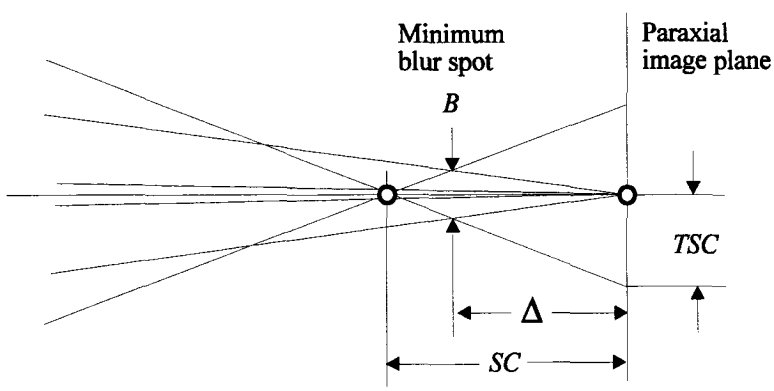


FIG. 3.11 Blur spot due to spherical aberration.

$B \approx 0.5\ TSC$, (*TSC* is the transverse spherical aberration contribution).
$\Delta \approx 0.75\ SC$, (*SC* is the longitudinal spherical aberration contribution).

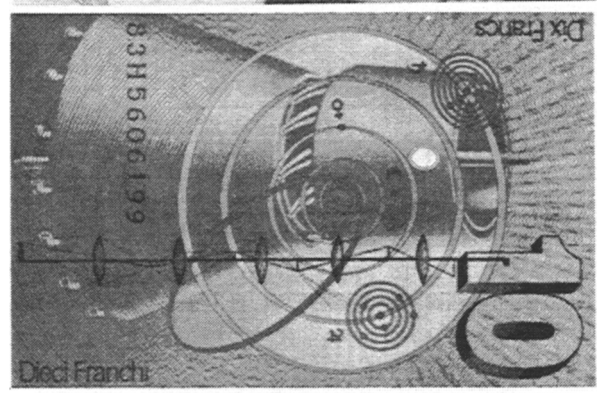
It is interesting to note that the relations expressed in Eqs. (3.5) and (3.6) were already stated in 1762 by the Swiss mathematician Leonard Euler.

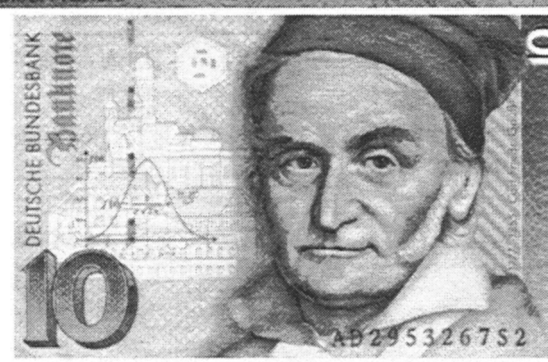
It is also interesting to observe that some countries honor their scientists by placing portraits of them on their bank notes with some remarks relating to the scientist's achievement.

Leonard Euler is pictured on the Swiss SFr 10 bank note with a train of lenses.

Karl Friedrich Gauss (1777–1855), the German mathematician and astronomer who was the first to use the paraxial-ray approximation, is pictured on the German DM 10 bank note with his distribution curve. Unfortunately, this will change when the Euro, the currency of the European Union, is introduced.

Equations (3.5)-(3.7) are applied in Table 3.1 for a few singlets made from the most common infrared materials and compared to a glass lens. For simplicity, the refractive indices shown are rounded off.





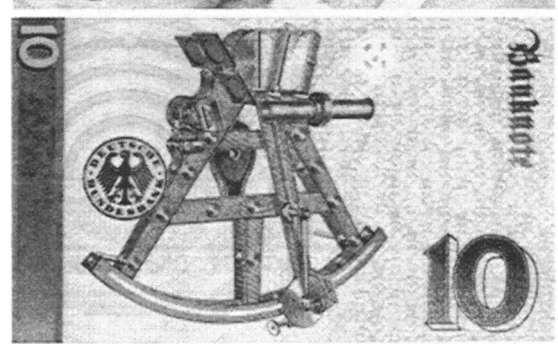


Table 3.1 Minimum spherical blur spot sizes for singlets made from different materials.

Material	N	R_1	R_2	β_{spher} (radians)
Glass	1.5	$0.583f$	$-3.5f$	$0.0670\,(f/\#)^{-3}$
Zinc Selenide	2.4	$0.885f$	$2.406f$	$0.0187\,(f/\#)^{-3}$
Silicon	3.4	$0.977f$	$1.649f$	$0.0108\,(f/\#)^{-3}$
Germanium	4.0	$1.0f$	$1.5f$	$0.0087\,(f/\#)^{-3}$

It is easy to remember that the best-shaped thin germanium lens, (shaped for minimum spherical aberration) has a front radius equal to its focal length and a rear radius that is 1.5 times its focal length (see also Fig. 3.10).

From the denominator of Eq. (3.6) it is apparent that for a convex plano lens ($R_2 = \infty$), $N(2N-1) = 4$. This means that the index for such a convex plano lens is $N = 1.686$. That is the index of sapphire at 3.7 μm. The front radius of such a lens would be $0.686f$ and the angular blur spot $\beta_{spher} = 0.0436\,(f/\#)^{-3}$ radians. Awareness of these simple relations can often save time, money, and disappointment.

We will now determine the remaining Seidel aberrations for a lens shaped for minimum spherical aberration.

3.3.2 Coma

As already mentioned and repeated in here Fig. 3.12, the third-order coma patch resembles the shape of a comet. Even though sagittal coma *CC* is only one third in size of the tangential coma, the triangle covered by the sagittal dimension contains about 55% of all the energy from the imaged object point.

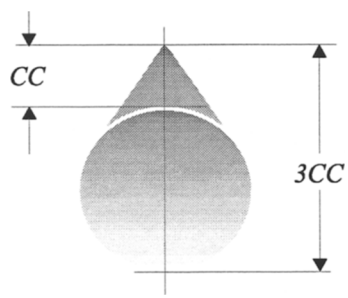


FIG. 3.12 Third-order coma patch.

If we use the amount of sagittal coma as a measure of the blur spot, its angular extent can be expressed by

$$\beta_{\text{coma}_{\text{min sph}}} = \frac{CC}{f} = \frac{u_p}{16(N+2)(f/\#)^2}, \tag{3.8}$$

where u_p is the angle of the principal ray for the specified off-axis point. Notice Eq. (3.8) is for a thin lens, shaped for minimum spherical aberration.

3.3.3 Astigmatism

The tangential and sagittal image shells are separated by $2AC = u_p^2 f$. The circle of least confusion is located halfway between (see Fig. 3.13) and has a diameter of

$$B = 2TAC = u_p^2 y. \tag{3.9}$$

Dividing this by the focal length of the lens, the angular blur spot due to astigmatism is

$$\beta_{\text{ast}} = \frac{B}{f} = \frac{u_p^2}{2(f/\#)}. \tag{3.10}$$

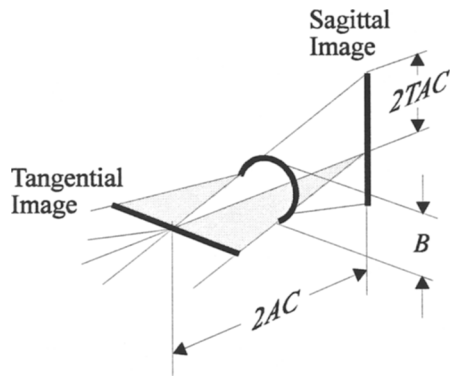


FIG. 3.13 Astigmatism, minimum blur spot, size, and location.

3.3.4 Field curvature

The Petzval surface (Fig. 3.14) is actually parabolic in shape and is expressed as Petzval contribution *PC* by

$$PC = -\frac{u_p^2 f}{2N}. \tag{3.11}$$

The vertex radius of this parabola is called the Petzval Radius ρ. Its measure for a thin lens is

$$\rho = -Nf. \tag{3.12}$$

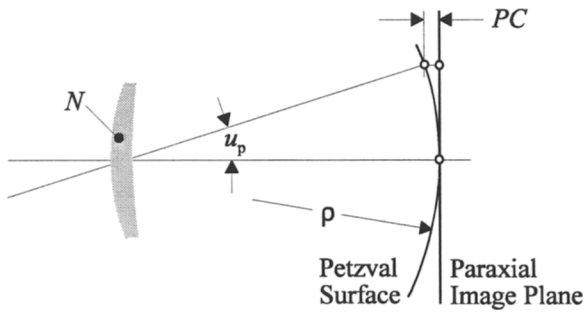


FIG. 3.14 Inward-curving Petzval surface for positive singlet.

3.3.5 Astigmatism and field curvature combined

Figure 3.15 indicates the relative locations among the astigmatism shells, the Petzval surface and the paraxial image plane at some off-axis point in the image. Notice that even if astigmatism is eliminated, which of course requires more than just a single element, field curvature will still exist. On the other hand, if the Petzval curvature were zero, i.e., coincident with the paraxial image plane, astigmatism would not be automatically corrected. It is important to understand the interrelation between astigmatism and field curvature to assess their contribution to image degradation.

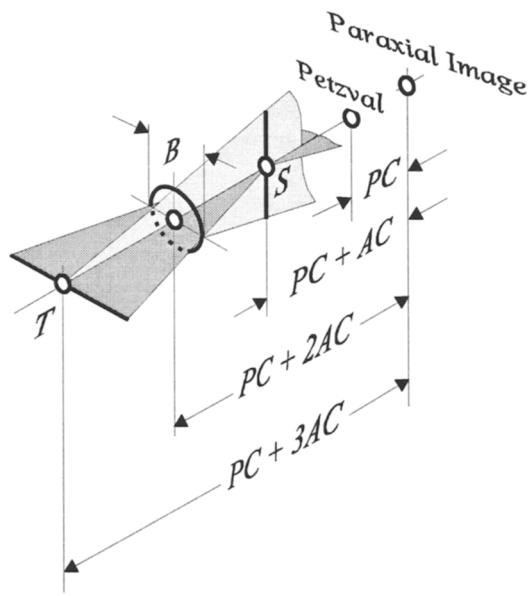


FIG. 3.15 Astigmatism and field curvature.

3.3.6 Axial chromatic aberration

As stated in Eq. (3.3), the power or inverse focal length of a thin lens is $\frac{1}{f} = (N-1)c$. Since indices of refraction vary with wavelengths, we can write for λ_1 and λ_2, respectively, $\frac{1}{f_1} = (N_1 - 1)c$ and $\frac{1}{f_2} = (N_2 - 1)c$. If these two wavelengths represent the limits of the spectral window under consideration, then the difference of the two focal lengths, $f_2 - f_1$ is the longitudinal axial chromatic aberration *LchC* as identified in Figs. 3.7 and 3.16. Forming this difference yields

$$LchC = f_2 - f_1 \cong \frac{N_1 - N_2}{N-1} f . \tag{3.13}$$

$\frac{N-1}{N_1 - N_2}$ is called inverse relative dispersion or Abbe number *V*, named after Ernst Abbe (1840–1905), chief optical designer and partner of Carl Zeiss, Germany. For the visible spectrum, the chosen wavelengths for the Abbe number are 0.5876 μm, 0.4861 μm, and 0.6563 μm. These are identified as Fraunhofer absorption lines d, F, and C (yellow, blue-green, and red). At the age of 27 years, Josef von Fraunhofer discovered and cataloged several hundred of these absorption lines from the sun spectrum. He classified the stronger absorbers with capital letters.[3]

With this, the Abbe number for the visible spectrum is stated by

$$\nu = \frac{N_d - 1}{N_F - N_C} . \tag{3.14}$$

For the infrared region, the Abbe number has been modified to read

$$\nu = \frac{N_M - 1}{N_S - N_L} , \tag{3.15}$$

where M, S, and L stand for middle, short, and long wavelengths.

Going back to Eq. (3.13), we can now write

$$LchC = \frac{f}{\nu} . \tag{3.16}$$

Figure 3.16 shows the relation between longitudinal axial chromatic aberration *LchC* and the blur spot diameter *B*, which is approximately equal to the transverse axial chromatic aberration *TAchC*. With $u' = 1/[2(f/\#)]$, the measure of the blur spot diameter becomes

$$B = LchC\, u' = \frac{LchC}{2(f/\#)} = \frac{f}{2v(f/\#)} \,. \tag{3.17}$$

The angular blur spot is then

$$\beta_{chrom} = \frac{B}{f} = \frac{1}{2v(f/\#)} \,. \tag{3.18}$$

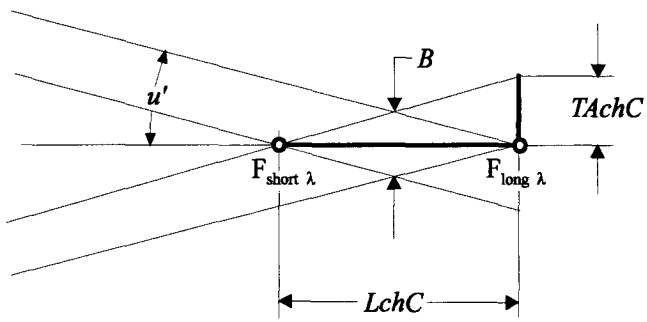


FIG. 3.16 Longitudinal and transverse axial chromatic aberrations (*LchC* and *TAchC* respectively). *B* is the minimum blur spot (circle of least confusion).

For the sake of completeness it must be mentioned that for a thin lens, with the stop at the lens, third-order lateral color and distortion are zero.

3.3.7 Numerical example

Let us look at a thin singlet made from silicon that is used in the 3- to 5-μm region. The chosen focal length is 100 mm, the lens diameter is 50 mm, and the half field to be covered is 0.1 rad (about 5.7 deg).

For the chosen spectral range, $N = N_M = 3.4254$, and $v = 235.5$. (These data can vary from one material supplier to another).

Applying the equations presented, beginning with Eq. (3.5), yields:

The radii for the lens with minimum spherical aberration are $R_1 = 97.9$ mm, and $R_2 = 164.1$ mm, i.e., a meniscus-shaped element. The Petzval radius $\rho = -342.5$ mm. The negative sign indicates an inward curved field, i.e., one concave in shape. The angular aberration blurs amount to

$$\begin{aligned} \beta_{spher} &= 1.33 \text{ mrad} \\ \beta_{coma} &= 0.29 \\ \beta_{astig} &= 2.50 \\ \beta_{chrom} &= 1.06 \\ \hline \Sigma\beta &= 5.18 \text{ mrad.} \end{aligned}$$

The linear blur spot size $B = f\Sigma\beta = 0.518$ mm.

For comparison and validation of the concept of third-order thin lens calculations, this lens was given a thickness of 5 mm and its shape was optimized for the given specification. The result is summarized in Fig. 3.17. Rays uniformly distributed across the entrance pupil form an image pattern (spot diagram) as they penetrate the image plane. The size and shape of the spot diagram indicates how well the lens can concentrate the transmitted energy, which is of course a measure of the aberrations present. To obtain the spot diagrams shown for our simple singlet, several hundred rays were traced for each object point. The surrounding box size in the diagrams was chosen to be 0.518 mm, the size of the third-order blur spot. We can see that the thin lens approximations compare very well indeed with the full scale geometric ray tracing results. To better demonstrate the presence of field curvature for a single lens element, spherical aberration was removed by aspherizing the front surface, and monochromatic radiation was assumed. Figure 3.18 shows the result; one can see distinctly the formation of the astigmatic image lines. Depending upon the application, one may want to "equalize" the spot size somewhat by refocusing the system as it was done for Fig. 3.17. The principle is illustrated in Fig. 3.19.

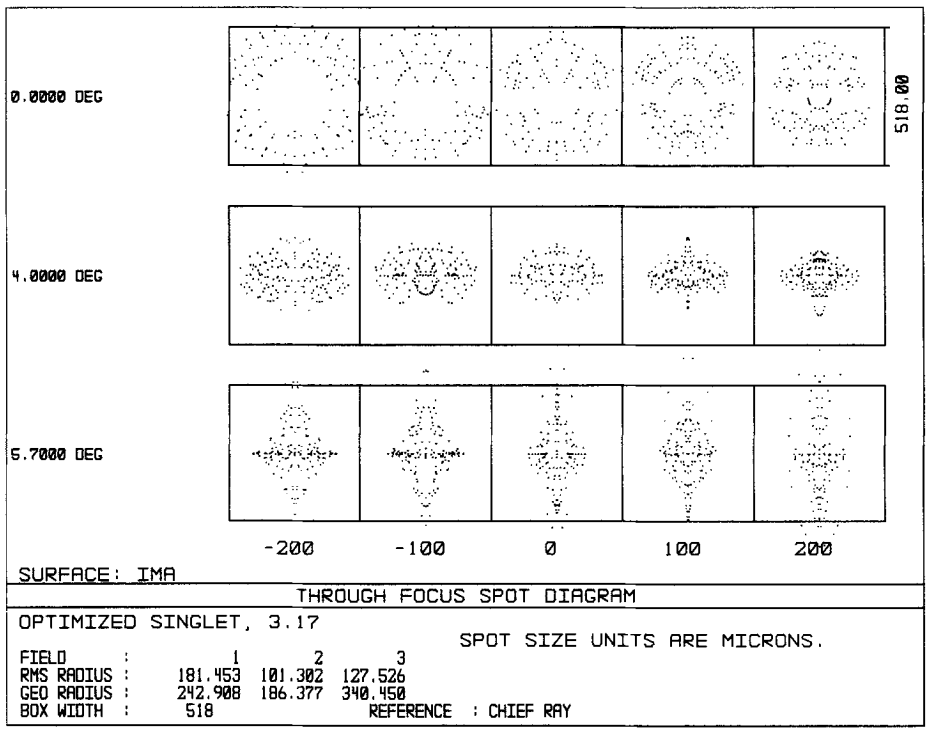


FIG. 3.17 Optimized singlet with added thickness.

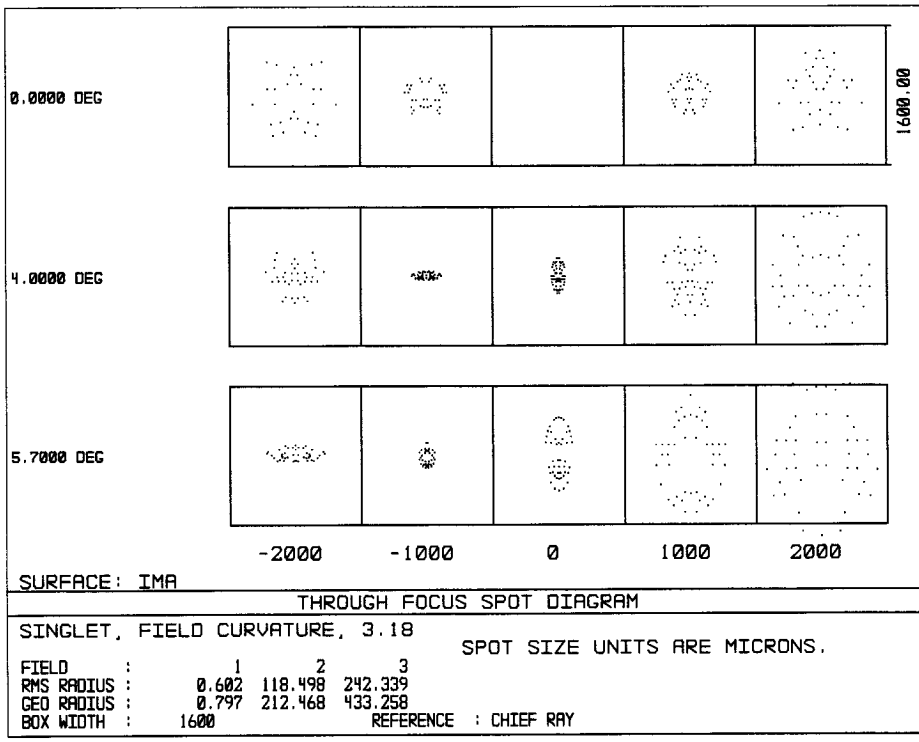


FIG. 3.18 Presence of field curvature and astigmatism for lens.

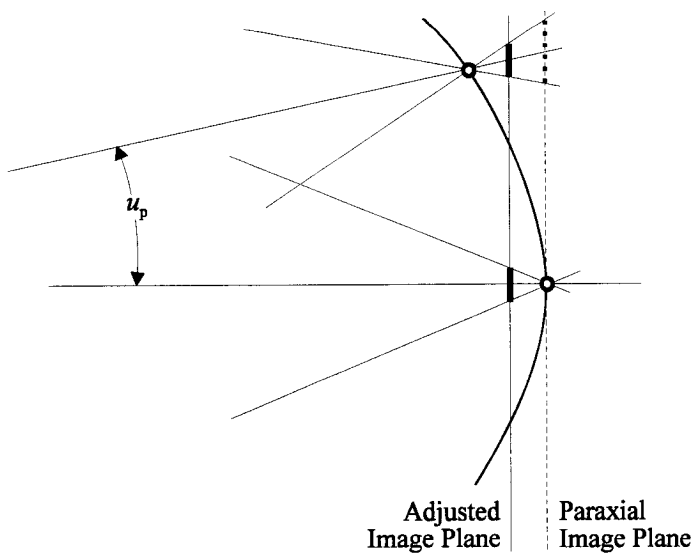


FIG. 3.19 Image position adjustment to compensate for blur spot size variation.

3.4 General Aberration Correction Methods

A very simple yet powerful approach to correcting off-axis aberrations without affecting spherical and axial chromatic aberrations is to place the aperture stop away from the lens, as shown in Fig. 3.20. The benefit of this method is achieved by controlling the size and location of the ray bundles forming the image. Shifting the stop a distance from the lens does, however, require an increase in the lens diameter; otherwise, vignetting occurs. Vignetting is the failure of an oblique ray bundle to fill the aperture stop. Relatively simple expressions have been developed to assess the effects of moving the aperture stop along the optical axis. These expressions are called *stop shift equations* and are treated in detail in Ref. 4.

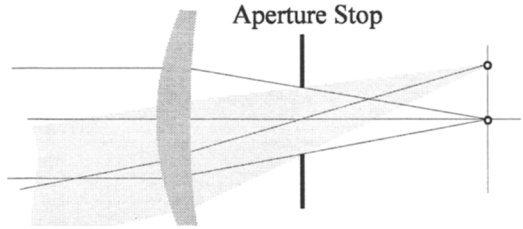


FIG. 3.20 Position and size of aperture stop control off-axis bundle.

Besides lens bending and proper positioning of the aperture stop, there are other general methods to be applied for aberration correction or balancing. Splitting a lens into more elements, aspherizing surfaces, carefully selecting the most suitable lens materials with regard to index and dispersion, combining refractive elements with reflective or diffractive components—all are techniques for improving image quality. Some configurations that employ these techniques will be discussed in a fundamental fashion, in line with the goal of this tutorial text.

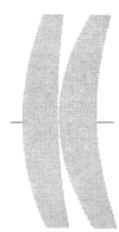
3.5 Doublets

Doublets are defined as two elements in close proximity. For the visible spectrum, doublets are frequently cemented together. Because optical cement is not transmissive in the infrared region, the elements are always separated from each other. The separation of the elements makes mounting somewhat more difficult, but one more degree of freedom is gained by being able to bend each element independently.

Let us look at some examples.

3.5.1 Two elements, same material

From the many combinations possible, we look in detail at the arrangement where each element is bent for minimum spherical aberration and made from the same material. As before, the object is considered to be located at infinity. There is no axial gap between the two thin elements and each element has half the power of the doublet. This means that the focal length of each element is twice as long as the focal length f of the doublet.



With these assumptions, and applying the thin lens third-order theory[4], the radii of the elements can be stated by

$$R_1 = \frac{4(N+2)(N-1)}{N(2N+1)} f, \tag{3.19}$$

$$R_2 = \frac{4(N+2)(N-1)}{N(2N-1)-4} f, \tag{3.20}$$

$$R_3 = \frac{4(N+2)(N-1)}{N(2N+1)+4(N^2-1)} f, \tag{3.21}$$

and

$$R_4 = \frac{4(N+2)(N-1)}{N(2N-1)+4(N^2-2)} f. \tag{3.22}$$

The angular blur spot size for this doublet configuration caused by spherical aberration is

$$\beta_{\text{spher}_{\min}} = \frac{N[5+4N(N-3)]}{512(N-1)^2(N+2)(f/\#)^3}. \tag{3.23}$$

Inspecting Eq. (3.23) reveals that the spherical aberration vanishes if N = 2.5. Figure 3.21 shows the dependency of β_{spher} on the material index N.

The angular coma blur spot size is

$$\beta_{\text{coma}} = \frac{u_p}{16(N+2)(f/\#)^2}. \tag{3.24}$$

The reversal at N = 2.5 indicates the change from undercorrected to overcorrected spherical aberration. Undercorrected means that the minimum blur spot lies inside the paraxial focal plane. For the overcorrected case, the location of the blur is outside the focal plane. This reversal indicates also that doublets made from material with an index of refraction greater than 2.5 can be corrected for spherical aberration by properly bending the elements.

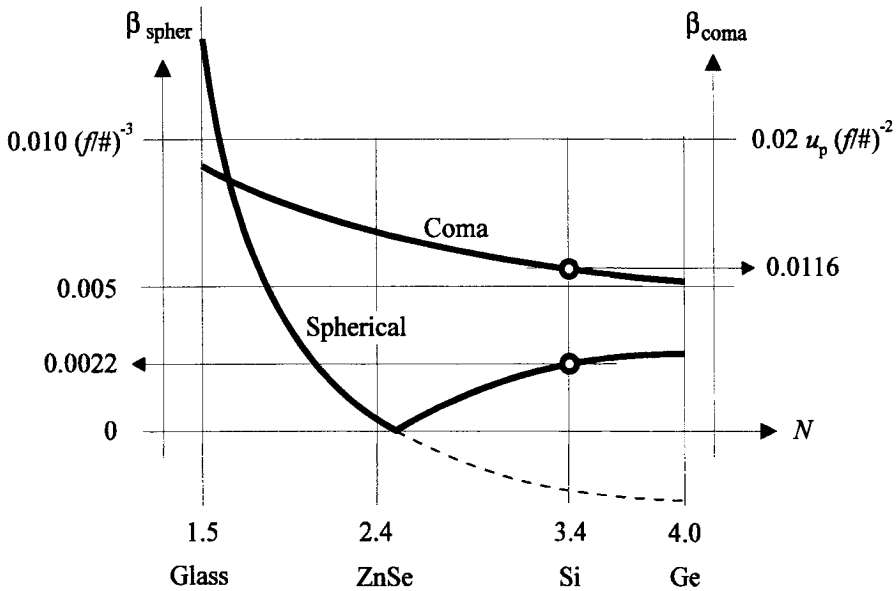


FIG. 3.21 Spherical aberration and coma as a function of the index. The example shows the coefficients for a silicon doublet. The aberrations are expressed in radians.

It is an educational exercise to develop expressions for a specific configuration. For example, two elements, equal in shape, bent as a combination for minimum spherical aberration may be of economic interest. Another doublet form would be two convex-plano elements. This is a particularly interesting exercise because it turns out that there is no gain with regard to spherical aberration, unless the material chosen has an index of refraction lower than 2. For IR applications, sapphire fits into this category.

3.5.2 Two elements, different materials

To find a reasonable starting point for a doublet with two different materials, we assume that the surface radius of the first element (element A) is equal but opposite in sign to the first surface radius. The first or front radius of the second element (element B) is equal to the second radius of element A. Finally, the rear surface of element B is a plano as shown in Fig. 3.22.

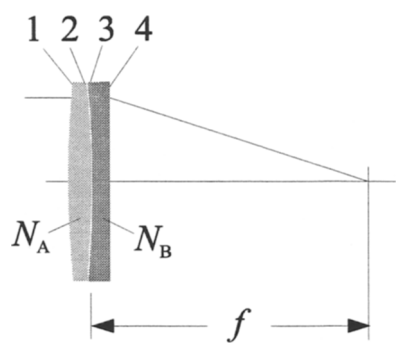


FIG. 3.22 Simple starting configuration for doublet with two different materials.

Expressing the above statement mathematically leads to $R_2 = R_3 = -R_1$, and $R_4 = \infty$. With that,

$$R_1 = (2N_A - N_B - 1)f\,. \tag{3.25}$$

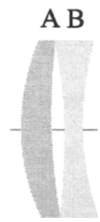
The focal length of the front element A is

$$f_A = \frac{R_1}{2(N_A - 1)} = \frac{(2N_A - N_B - 1)}{2(N_A - 1)} f\,, \tag{3.26}$$

and for the rear element B, the focal length is

$$f_B = -\frac{R_1}{(N_B - 1)} = -\frac{(2N_A - N_B - 1)}{(N_B - 1)} f\,. \tag{3.27}$$

3.5.3 The achromat



A thin achromat can be corrected for third-order spherical aberration, coma, and axial chromatic aberration. The combination consists of a positive and a negative element. To eliminate axial chromatic aberration, a simple relation between powers and Abbe numbers of the two elements must be met. This relation is

$$\phi_A = \frac{\nu_A}{(\nu_A - \nu_B)}\phi. \tag{3.28}$$

And, since $\phi = \phi_A + \phi_B$,

$$\phi_B = \frac{\nu_B}{(\nu_B - \nu_A)}\phi\,. \tag{3.29}$$

The process of deriving the complete prescription for such a doublet is straightforward but too cumbersome to be included here.[5]

The radii as a factor of an achromat's focal length are listed in Table 3.2 for two objectives: one for the MWIR and one for the LWIR region. For the MWIR lens, silicon and germanium have been the materials chosen, and the combination Amtir-1/zinc sulfide was selected for the LWIR objective. Amtir is an acronym for amorphous material transmitting infrared radiation. The composition of Amtir-1 is 33% Ge, 12% As, and 55% Se.

Table 3.2 Radii of lens elements for two selected achromats.

Spectral region	**MWIR** (3 – 5 μm)	**LWIR** (8 – 12 μm)
Front element	Silicon	Amtir-1
R_1	$0.97f$	$1f$
R_2	$3.25f$	$6f$
Rear element	Germanium	Zinc sulfide
R_3	$4f$	$-6f$
R_4	$2f$	$-24f$

These choices are a sound starting point that lead quickly to good solutions in terms of optimization with the computer after adding thicknesses and spacings.

3.6 Two Thin Air-Spaced Elements

Separating the two elements provides an additional degree of freedom that can be applied for aberration correction or to determine the basic configuration of the system. The Petzval objective, telephoto lenses, telescopes, and eyepieces are all examples that belong to this category.

Much can be learned and a good starting point can be obtained by modeling these arrangements with separated thin lenses.

With a few judgments and assumptions, closed-form solutions for the elements' shape can be derived for some configurations. Simple expressions provide a quick answer in the layout stages for trade-off evaluation among different system candidates.

3.6.1 The Petzval objective

As shown in Fig. 3.23, this lens consists in principle of two elements, spaced by approximately the focal length f of the lens.

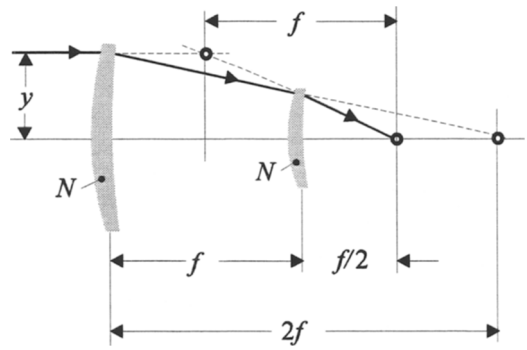


FIG. 3.23 Basic configuration of the Petzval objective.

The focal length of the front element is twice that of the lens; the focal length of the rear element is equal to the focal length of the lens. We further stipulate that the elements are individually shaped for minimum spherical aberration and made from the same material. Again, applying the thin lens third-order aberration equations, the four radii for this objective can be derived:

$$R_1 = \frac{4(N+2)(N-1)}{(2N+1)N} f, \tag{3.30}$$

$$R_2 = \frac{4(N+2)(N-1)}{N(2N-1)-4} f, \tag{3.31}$$

$$R_3 = \frac{2(N+2)(N-1)}{N(6N+1)-4} f, \tag{3.32}$$

and

$$R_4 = \frac{2(N+2)(N-1)}{N(6N-1)-8} f. \tag{3.33}$$

The angular blur spot due to spherical aberration is

$$\beta_{\text{spher}} = \frac{N\left[4N(2N-7)+11\right]}{2048(N-1)^2(N+2)(f/\#)^3}. \tag{3.34}$$

Axial chromatic aberration amounts to

$$\beta_{\text{chrom}} = \frac{3}{8v(f/\#)}. \tag{3.35}$$

Based on these equations, a germanium objective with 100 mm focal length and a relative aperture of $f/1$, used for the 8- to 12-μm region, would yield an on-axis blur spot size diameter of 0.135 mm. A very good starting point indeed, if one considers that this is

only several times the size of the diffraction limit. In Chapter 6 it will be shown how to reduce the spherical aberration by changing one surface from spherical to aspherical, and by adding a diffractive phase profile to one surface to lessen the chromatic aberrations of this objective.

Closed-form equations can be also developed for the off-axis aberrations by applying the same general procedure[4] that was used to derive Eqs. (3.34) and (3.35). However, they become very lengthy and are no longer simple expressions. Therefore, they will not be included.

3.6.2 Refractive beam expanders

The two basic configurations used for these afocal systems are the Galilean and the Keplerian telescope types, shown in Figs. 3.24 and 3.25.

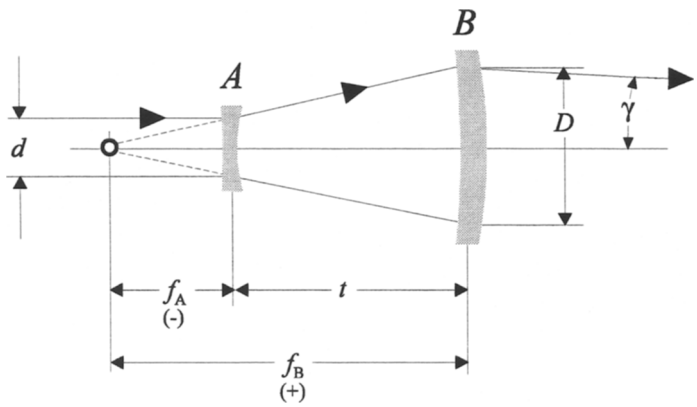


FIG. 3.24 Galilean beam expander.

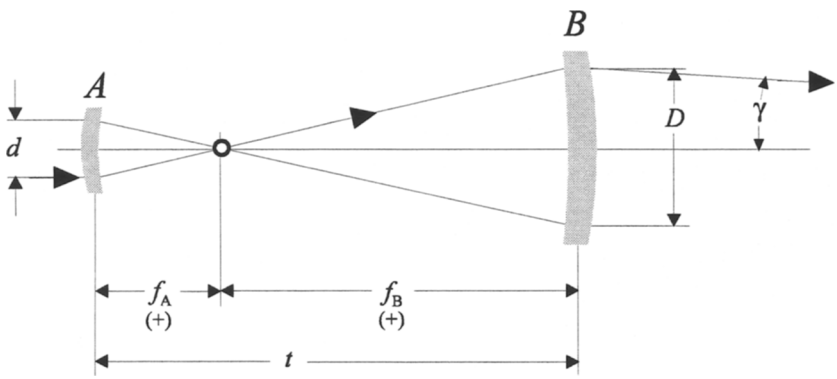


FIG. 3.25 Keplerian beam expander.

Since each of the two elements is suffering from spherical aberration, the exiting beam is not precisely collimated. The marginal ray converges under an angle γ. If each element is shaped individually for minimum spherical aberration, this convergence angle is

$$\gamma = \frac{N(4N-1)}{64(N-1)^2(N+2)(f/\#)^3}\left[\frac{m-1}{m}\right]. \tag{3.36}$$

N is the index of refraction (both elements are from the same material). The magnification of the system m is expressed by the ratio of the exiting and entering beam diameters D and d. In Eq. (3.36), m is positive for the Galilean system. Because of the image reversal, it is negative for the Keplerian configuration. It is apparent that the convergence for the Galilean system is less than for the Keplerian, because the negative first element is afflicted with over-corrected spherical aberration, compensating somewhat for the larger amount of under-corrected spherical aberration from the positive second element. It will be noticed that the factor before the square bracket in Eq. (3.36) is twice the value of the minimum angular blur spot size for a single thin lens due to spherical aberration. The relative aperture $f/\# = f_B/D = |f_A|/D$, and the separation of the elements $t = f_A + f_B$. For a germanium beam expander with $N = 4, f/\# = 2$, $m = 5$, the convergence angles are $\gamma_{Gal} = -1.7$ mrad, and $\gamma_{Kep} = -2.6$ mrad.

The fact that a negative lens has over-corrected spherical aberration, as mentioned above, permits us to bend the negative entrance element of a Galilean beam expander in such a way that it will compensate completely for the under-corrected spherical aberration of the positive exit lens.

If the positive element B is shaped for minimum spherical aberration, its radii are

$$R_{B1} = -\frac{2(N_B+2)(N_B-1)}{N_B(2N_B-1)-4} f_B, \text{ and} \tag{3.37}$$

$$R_{B2} = -\frac{2(N_B+2)(N_B-1)}{N_B(2N_B+1)} f_B. \tag{3.38}$$

To compensate for the spherical aberration of this element, the proper shape of the negative element A can be extracted from the quadratic equation

$$\frac{(N_A+2)}{N_A} K_A^2 - (2N_A+1)K_A + N_A^2 - \frac{N_B(4N_B-1)(N_A-1)^2 m}{4(N_B-1)^2(N_B+2)} = 0. \tag{3.39}$$

m is the expansion ratio or magnification of the beam expander. It is expressed by

$$m = -\frac{f_B}{f_A} = \frac{D}{d}, \tag{3.40}$$

where D and d are the free aperture diameters of the elements, as indicated in Fig. 3.24.

K_A is the lens shape factor of element A. Rearranging an earlier given expression for the shape factor yields for the first radius of element A

$$R_{A1} = \frac{(N_A - 1)}{K_A} f_A, \text{ and} \tag{3.41}$$

$$R_{A2} = \frac{(N_A - 1)}{(K_A - 1)} f_A. \tag{3.42}$$

For the case when both elements are made from the same material, Eq.(3.39) reduces to

$$\frac{(N+2)}{N} K_A^2 - (2N+1)K_A + N^2 - \frac{N(4N-1)m}{4(N+2)} = 0. \tag{3.43}$$

Table 3.3 lists the radii of four Galilean beam expanders using different materials for the elements. The focal length of element B was chosen to be 50 mm for all cases. The magnification of the systems is 5.

Table 3.3 Radii in millimeters of Galilean beam expanders, corrected for spherical aberration.

Material	N	K_A	R_{A1}	R_{A2}	R_{B1}	R_{B2}
Glass	1.5	1.81547	−2.750	−6.131	+175.00	−29.15
Zinc Selenide	2.4	3.18146	−4.401	−6.412	−120.30	−44.25
Silicon	3.4	4.69063	−5.117	−6.503	−82.45	−48.85
Germanium	4.0	5.58199	−5.374	−6.548	−75.00	−50.00

Optimizing the thin lens values with added thicknesses of 0.8 and 1.2 mm results in the arrangement shown in Fig. 3.26. In optimization, only the second and forth radii were modified in addition to the adjusted spacing between the elements. While the process was executed for a single wavelength—namely 10.6 μm—the CO_2 laser wavelength, this beam expander performs well over the entire LWIR band from 8 to 12 μm.

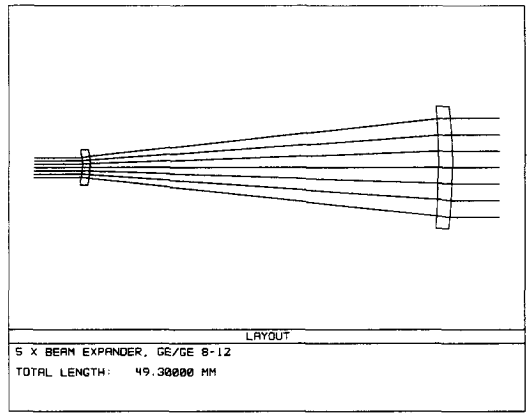


FIG. 3. 26 Galilean 5× beam expander.

3.6.3 Telephotos

Telephotos and reversed telephoto systems can be analyzed the same way with the thin lens concept for predesign purposes. The two types are shown in Fig. 3.27.

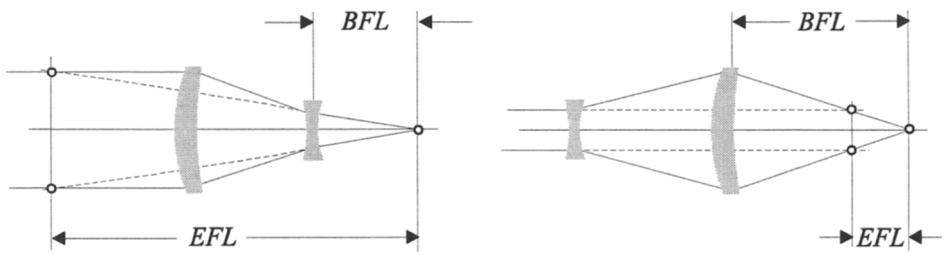


FIG. 3.27 Telephoto and reversed telephoto lenses.

A telephoto lens is compact for its long focal length. A reversed telephoto has a short focal length compared to its overall size. Notice the striking difference in the ratios between the effective focal lengths *EFL* (system focal length *f*) and the back focal lengths *BFL* (the distance from the last optical surface to the focal plane).

3.7 Reflective Optics

There are several advantages to using reflective optical elements; one is that mirrors are not wavelength selective. This is very convenient for system alignment; the user can focus the system in the visible spectrum and then place the detector at the same image position. Anyone who has had to align an IR system "in the dark," i.e., with a detector alone, knows what a relief it is to be able to observe visually what is going on. Imaging mirrors are usually less expensive than lenses made for the infrared because most of the time the mirror substrate material is of secondary importance. The demand is more of a mechanical nature. However, there are also disadvantages. One of them is that the image formed by a mirror lies at the same side as the object. This means that the detector is in the way, resulting in the obstruction of incoming radiation as it is shown in Fig. 3.28. As

we shall see shortly, there are methods to avoid the obstruction, but, as expected, one has to pay a price.

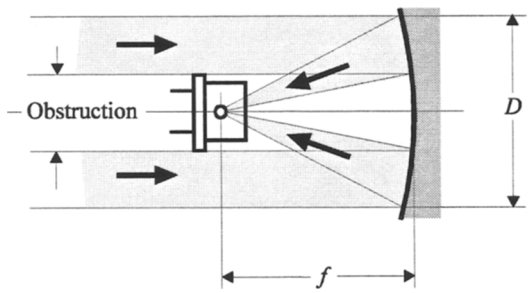


FIG. 3.28 Centered imaging mirror, showing central obstruction.

3.7.1 The spherical mirror

A mirror has only one optical surface. If the surface is a section of a sphere, it is relatively easy to manufacture and performs rather well. Spherical aberration of a spherical mirror with the same *f*–number as a glass lens, shaped for minimum spherical aberration, is approximately eight times smaller than that of the lens.

The angular blur spot due to third-order spherical aberration for a spherical mirror, with an object located at infinity and the aperture stop at the mirror, is

$$\beta_{\text{spher}} = \frac{1}{128(f/\#)^3} . \tag{3.44}$$

The coma blur is larger by a factor of $(N+2)$ than that for a best-shaped single lens with index N. It is

$$\beta_{\text{coma}} = \frac{u_p}{16(f/\#)^2} . \tag{3.45}$$

The angular astigmatic blur spot size is the same as for the thin lens, namely

$$\beta_{\text{astig}} = \frac{u_p^2}{2(f/\#)} . \tag{3.46}$$

Using a spherical mirror "off-axis" to avoid obstruction can be looked upon in several ways. This is indicated in Figs. 3.29, 3.30, and 3.31. Figure 3.29 shows the aperture stop decentered to select a portion of the total mirror. In Fig 3.30 the spherical mirror is laterally displaced from the optical axis and tilted. The aperture stop is the physical size of the mirror. Figure 3.31 has the mirror centered about the optical axis and the incoming collimated beam is oblique relative to the optical axis. The performance is the same for all three configurations because a section of a sphere remains a spherical surface of rotation about any normal of the surface.

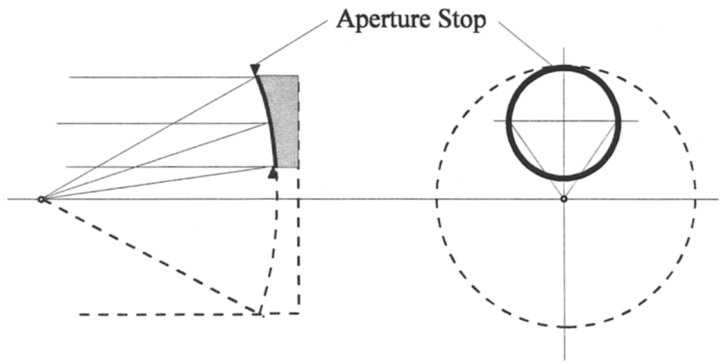


FIG. 3.29 Spherical mirror with decentered aperture stop.

FIG. 3.30 Tilted mirror, laterally displaced.

FIG. 3.31 Centered mirror with obliquely entering energy bundle.

Fig. 3.31 is presented three-dimensionally in Fig. 3.32 to illustrate how the sagittal and tangential foci separate with the change of u_p, the angle of the incoming collimated beam. While the sagittal focus F_S moves along a line as a function of angle u_p, the tangential F_T follows the path of a circle whose radius ρ is one quarter of the mirror radius of curvature.

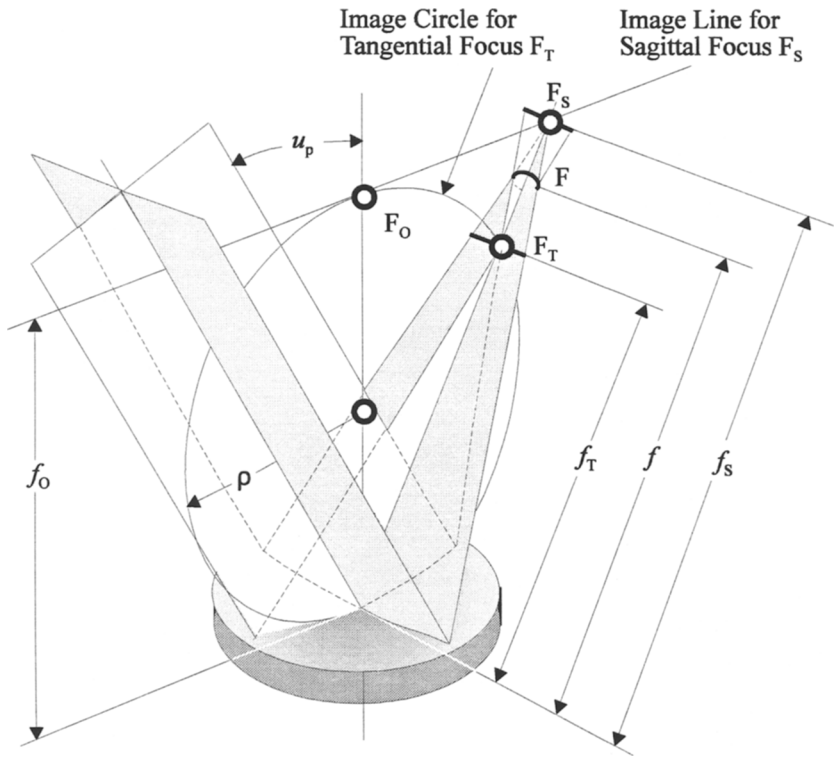


FIG. 3.32 Astigmatism of spherical mirror for an object at infinity.[6]

From Fig. 3.32 it can be reasoned that the sagittal focus can be made to coincide with the tangential focus if the mirror radius that focuses the sagittal ray is made shorter. A mirror with such a shape is a toroid. For a given focal length, the two radii are

tangential radius $$R_T = \frac{2f}{\cos u_p}, \tag{3.47}$$

sagittal radius $$R_S = 2f \cos u_p. \tag{3.48}$$

3.7.2 The Mangin mirror

This element is a thick negative meniscus with two spherical surfaces. The rear surface is the reflector A. Mangin, a French officer, invented it in 1876 to avoid the cost of generating a parabola. The overcorrected spherical aberration of the negative lens portion compensates for the undercorrected spherical aberration from the mirror. Since a ray passes twice through the meniscus (Fig. 3.33), the Mangin mirror functions like a triplet.

An element that uses refractive as well as reflective power, as is the case for the Mangin mirror, is called a catadioptric element.

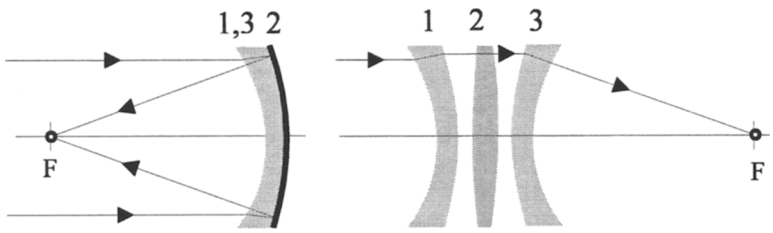


FIG. 3.33 The Mangin mirror and its equivalent lens triplet configuration.

Mangin elected to make the front surface concentric to the element's focus, where he placed the source for the search light.

The power for a thin Mangin mirror can be expressed by

$$\phi = \frac{1}{f} = \frac{2(N-1)}{R_1} - \frac{2N}{R_2}. \tag{3.49}$$

$$\text{If } R_1 = -f,\ R_2 = -\frac{2N}{(2N-1)}f, \tag{3.50}$$

and the aberrations for such a configuration are

spherical aberration $$\beta_{\text{spher}} = \frac{2N-3}{128N^2(f/\#)^3}, \tag{3.51}$$

coma $$\beta_{\text{coma}} = \frac{u_p}{16N^2(f/\#)^2}, \tag{3.52}$$

astigmatism $$\beta_{\text{astig}} = \frac{u_p^2}{2(f/\#)}, \text{ and} \tag{3.53}$$

axial chromatic aberration $$\beta_{\text{chrom}} = \frac{N-1}{2N\nu(f/\#)}. \tag{3.54}$$

Eq. (3.51) reveals that for the chosen arrangement ($R_1 = -f$), spherical aberration vanishes if $N = 1.5$. Third-order spherical aberration can be eliminated for other materials by applying the lens bending technique, which leads to solving the cubic equation[7]

$$\left(\frac{f}{R_1}\right)^3 + \frac{(N+3)}{2}\left(\frac{f}{R_1}\right)^2 + \frac{(4N+5)}{4}\left(\frac{f}{R_1}\right) + \frac{(4N^2-3)}{8(N-1)} = 0. \tag{3.55}$$

The radius for the reflective rear surface can be calculated by rearranging Eq. (3.49) to read

$$R_2 = \frac{2NR_1 f}{2(N-1)f - R_1}. \tag{3.56}$$

Table 3.4 lists the radii for Mangin mirrors made from different materials, corrected for third-order spherical aberration.

Table 3.4 Surface radii for Mangin mirrors, corrected for third-order spherical aberration.

Material	N	R_1	R_2
Glass	1.5	$-1.0f$	$-1.5f$
Zinc Selenide	2.4	$-1.1506f$	$-1.3980f$
Silicon	3.4	$-1.1790f$	$-1.3409f$
Germanium	4.0	$-1.1823f$	$-1.3169f$

3.7.3 Classical two-mirror configurations

There are four basic two-mirror configurations we shall discuss in detail. They are named after their inventors. All suffer from central obscuration. The systems are illustrated in Fig. 3.34.

1. The *Newton* configuration consists of a focusing primary mirror and a secondary flat folding mirror that deviates the light bundle by 90° relative to the optical axis to locate the focus of the arrangement outside the incoming beam.

2. The *Gregory* arrangement has two concave mirrors, opposing each other. An intermediate image is formed between the two mirrors. The final image is right side up, and therefore the configuration is a terrestrial telescope.

3. *Cassegrain* changed Gregory's setup somewhat by turning the secondary mirror around, making it a convex element. The system becomes much more compact. It is the most widely used configuration for infrared applications. The fact that the image is upside down is of not much concern. Such a type is referred to as an astronomical telescope. Newton never credited Cassegrain for inventing this compact system. He claimed that it was just a logical extension of Gregory's arrangement.

4. *Schwarzschild* reversed the basic Cassegrain type, making his configuration similar to a reversed telephoto. This objective is most frequently used as a microscope objective.

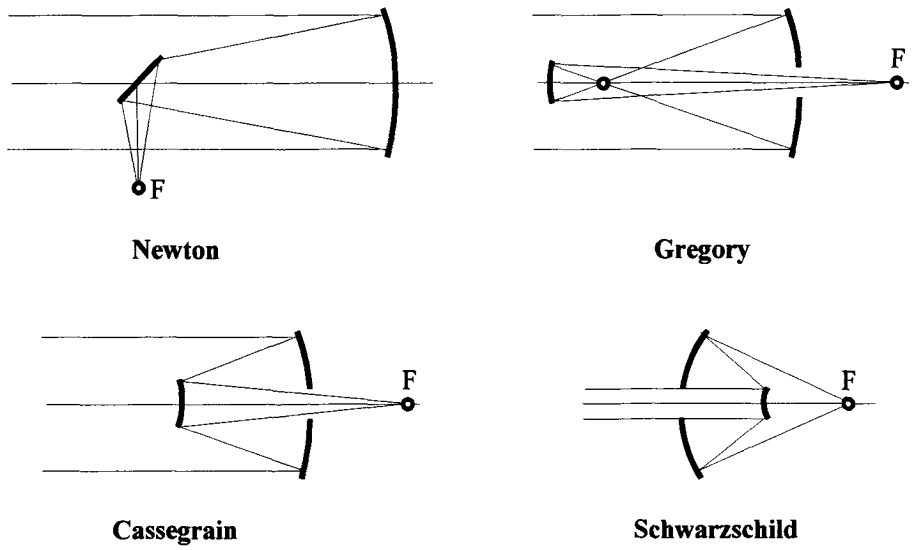


FIG. 3.34 The four classical two-mirror configurations.

3.7.4 The two-sphere Cassegrain system

As stated, because of its compactness, the Cassegrain configuration is frequently used for IR applications. Therefore, it is worthwhile to investigate what two spherical mirrors can do for this configuration. Figure 3.35 shows the physical length of the Cassegrain objective in relation to its focal length.

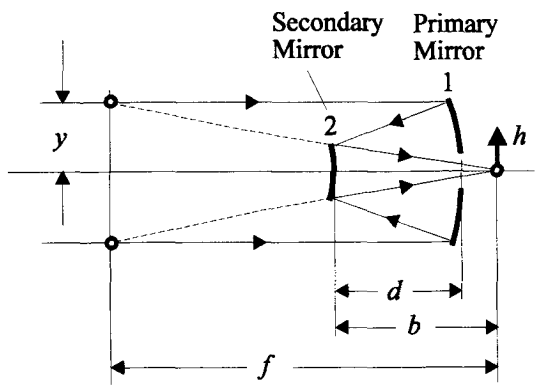


FIG. 3.35 Basic Cassegrain configuration dimensions. f is the effective focal length, d is the spacing between the two mirrors, b is the distance from the secondary mirror 2 to the image plane, called the back focal distance, y is the semi aperture size, and h is the image height. All dimensions identified are considered to be positive.

The mirror radii for this arrangement are

$$R_1 = \frac{2fd}{b-f}, \tag{3.57}$$

and

$$R_2 = \frac{2bd}{b+d-f}. \tag{3.58}$$

With the aperture stop at the primary mirror, the general third-order aberration contribution equations[8] yield the following expressions for the angular blur spots in radians:

Spherical aberration $$\beta_{\text{spher}} = \frac{f(b-f)^3 + b(f-d-b)(f+d-b)^2}{128fd^3(f/\#)^3}. \tag{3.59}$$

Coma $$\beta_{\text{coma}} = \frac{h\left[2f(b-f)^2 + (f-d-b)(f+d-b)(d-f-b)\right]}{32d^2f^2(f/\#)^2}. \tag{3.60}$$

Astigmatism $$\beta_{\text{astig}} = \frac{h^2\left[4bf(b-f) + (f-d-b)(d-f-b)\right]}{8f^3bd(f/\#)^3}. \tag{3.61}$$

The Petzval radius is $$\rho_{\text{Petzval}} = \frac{bfd}{fd-(b-f)^2}. \tag{3.62}$$

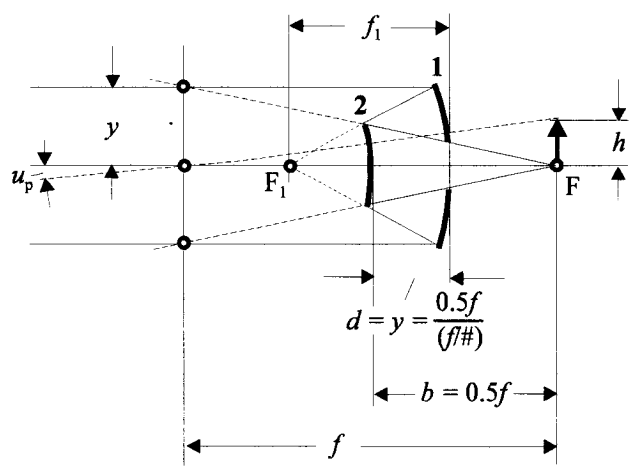


FIG. 3.36 Cassegrain system layout with 25% area obstruction and a relative aperture of $f/1$ for the primary mirror. The system's $f/\# = f/(2y)$, and the half-field angle $u_p = h/f$ for an object located at infinity.

With (f /1) as practical choice for the f-number of the primary mirror, and a 25% obstruction for the incoming energy bundle, the general layout from Fig. 3.35 changes to the one shown in Fig. 3.36.

The surface radii for the arrangement in Fig. 3.36 are

$$R_1 = -\frac{2f}{(f/\#)}, \tag{3.63}$$

and

$$R_2 = \frac{f}{1-(f/\#)}. \tag{3.64}$$

The angular blur spot sizes are

Spherical aberration

$$\beta_{\text{spher}} = \frac{1}{256}\left[1-(f/\#)^{-1}+(f/\#)^{-2}+(f/\#)^{-3}\right]. \tag{3.65}$$

Coma

$$\beta_{\text{coma}} = \frac{u_p}{64}\left\{4+\left[(f/\#)^2-1\right]\left[1-3(f/\#)\right](f/\#)^{-3}\right\}. \tag{3.66}$$

Astigmatism

$$\beta_{\text{astig}} = \frac{u_p^2}{16}\left\{\left[(f/\#)-1\right]\left[1-3(f/\#)\right]^2(f/\#)^{-3}-8\right\}, \tag{3.67}$$

and the radius of the Petzval curvature is

$$\rho = \frac{f}{2-(f/\#)}. \tag{3.68}$$

Spherical aberration reaches a minimum with f /3 where it is 0.003183 rad. Coma for that f-number is 0.025463 u_p, and astigmatism is 0.203704 u_p^2. The Petzval radius is $-f$, i.e., concave. If we assume a half-field angle of 0.026 rad (1.5°), and a 100-mm focal length f, the radii become $R_1 = -66.6667$ mm, and $R_2 = -50$ mm. The total angular blur spot size adds up to ≅ 4 mrad. The physical diameter is $0.004 \times 100 = 0.4$ mm. While this is relatively large, it may be adequate for the application under consideration.

To demonstrate the usefulness of the third-order calculations once more, a plot of the encircled energy distributions obtained by geometrical ray tracing is presented in Fig. 3.37 for comparison.

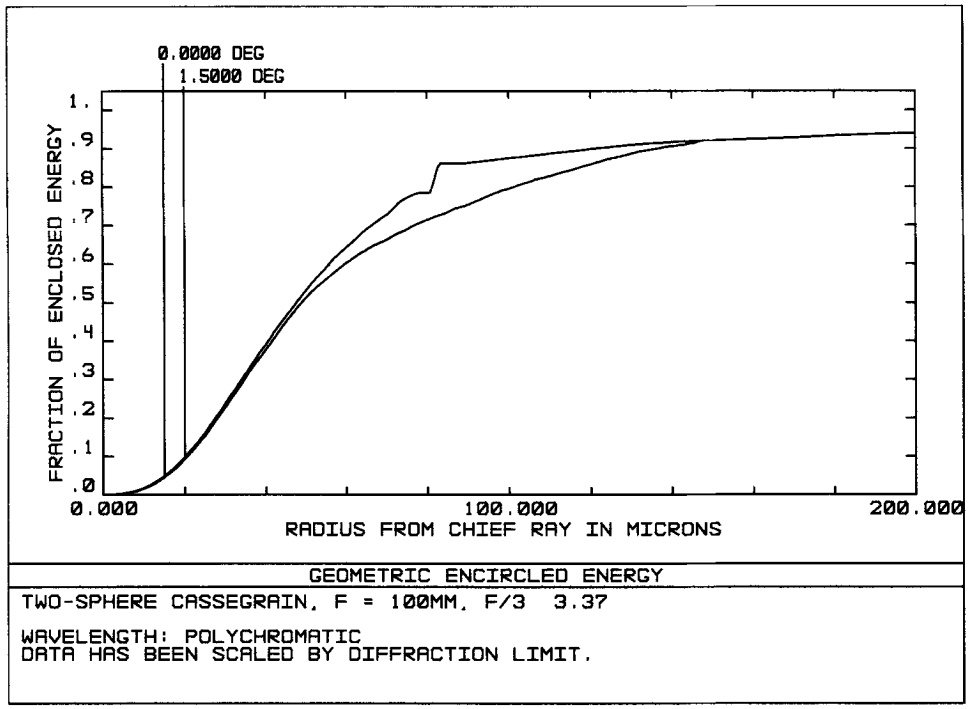


FIG. 3.37 Encircled energy plot obtained by geometric ray tracing.

3.7.5 The two-sphere Gregory system

In Fig. 3.38, the Gregory system is compared to the Cassegrain arrangement. For the same obstruction ratio between the secondary and primary mirrors and the same image position, the focal length and therefore the *f*-number increases for the Gregory system, and so does the overall physical length.

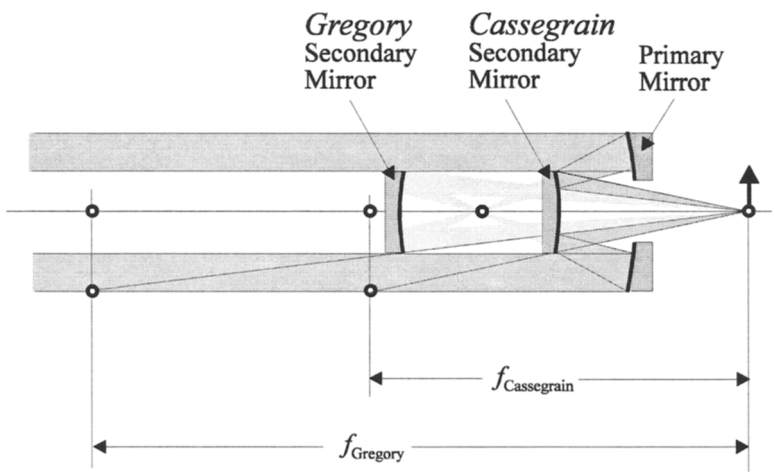


FIG. 3.38 Layout comparison between the Gregory and Cassegrain arrangements.

To calculate the radii of the spherical mirrors for the Gregorian system and its aberrations, Eqs. (3.57) through (3.62) can be applied by changing the focal length to a negative value. Contrary to the Cassegrain, where the mirrors have opposite shapes (concave primary and convex secondary), both mirrors for the Gregorian are concave, which results in a higher total spherical aberration. To improve the performance of both configurations, aspheric surfaces can be used. This will be discussed in Chapter 5.

3.7.6 Schwarzschild, a very special case

If the two spherical mirrors are concentric to each other, separated by twice the system's focal length, third-order spherical aberration, coma and astigmatism are eliminated, provided the aperture stop is located at the common center of curvature c as shown in Fig. 3.39 Adding a thin field flattener lens in the focal plane corrects for distortion and field curvature as well.

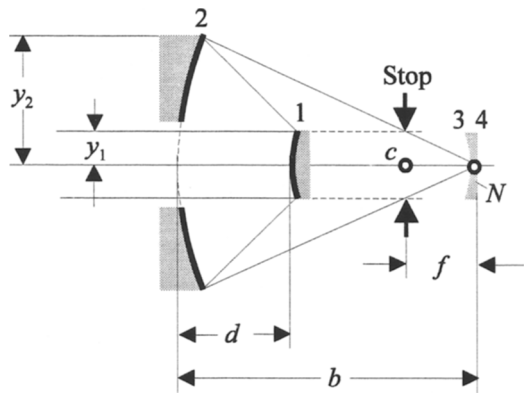


FIG. 3.39 The concentric Schwarzschild objective with field flattener.

The specific relations to achieve the corrections of spherical aberration, coma, and astigmatism are

$$
\begin{aligned}
d &= 2f, \\
b &= (\sqrt{5}+2)f, \\
R_1 &= (\sqrt{5}-1)f, \\
R_2 &= (\sqrt{5}+1)f, \\
y_2 &= (\sqrt{5}+2)y_1 .
\end{aligned}
\tag{3.69}
$$

The Petzval radius is, according to Eq. (3.62), $\rho = -f$. To correct for this effect, we insert a field lens whose Petzval radius is $+f$. Since the function of this field lens is to

"flatten" the image in the field, it is called a field flattener. By applying Eq. (3.12) to calculate the focal length of the field flattener, we obtain

$$f_{\text{flat}} = -\frac{f}{N}, \tag{3.70}$$

where N is the index of refraction for the field flattener lens, and f is the focal length of the Schwarzschild objective. By placing the field flattener in the plane of the image, the lens can be shaped as stated in Eqs. (3.71) and (3.72) to eliminate distortion.[9]

The front radius $$R_3 = \frac{N^2-1}{1-N(N+1)} f, \tag{3.71}$$

and the rear radius $$R_4 = (N^2-1) f. \tag{3.72}$$

With a high-index material, such as $N = 4$ for germanium, R_4 becomes 15 times the focal length. This suggests substitution of a flat surface and correction of R_3 for power. Not only would this lower the cost of the lens, but it invites the idea of depositing the detector elements directly onto the rear surface of the field flattener.

For further study, including the aberrations for finite conjugate configurations, see Ref. 10.

3.7.7 Reflective beam expanders

Reflective beam expanders are modified Cassegrain and Gregorian mirror systems. Both the object and the image are located at infinity, which means that the input and output beams are collimated (afocal system). We shall assess how well that can be achieved with two spherical mirrors. Figure 3.40 shows the Cassegrain configuration. To avoid obstruction, the mirrors are used in an off–axis mode. Due to the spherical shape of the mirrors, the exiting beam is not collimated. The marginal ray converges under an angle γ relative to the optical axis. Due to the off-set of the mirrors from the optical axis, the convergence angle α of the total beam in the meridional plane shown is somewhat smaller than γ. The Gregorian beam expander is shown in Fig. 3.41.

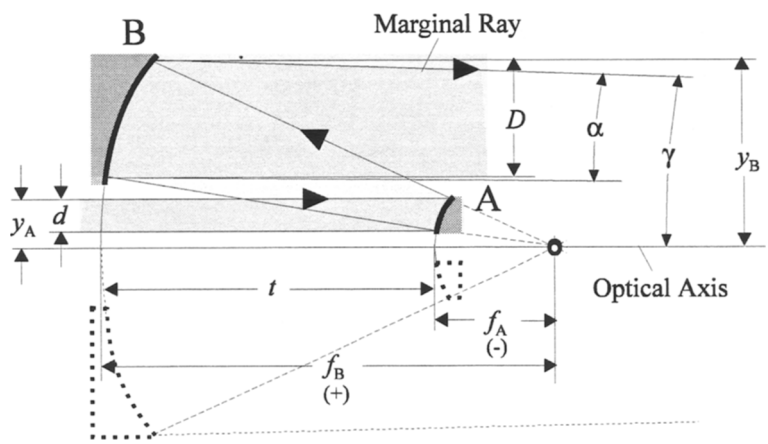


FIG. 3.40 Cassegrain beam expander.

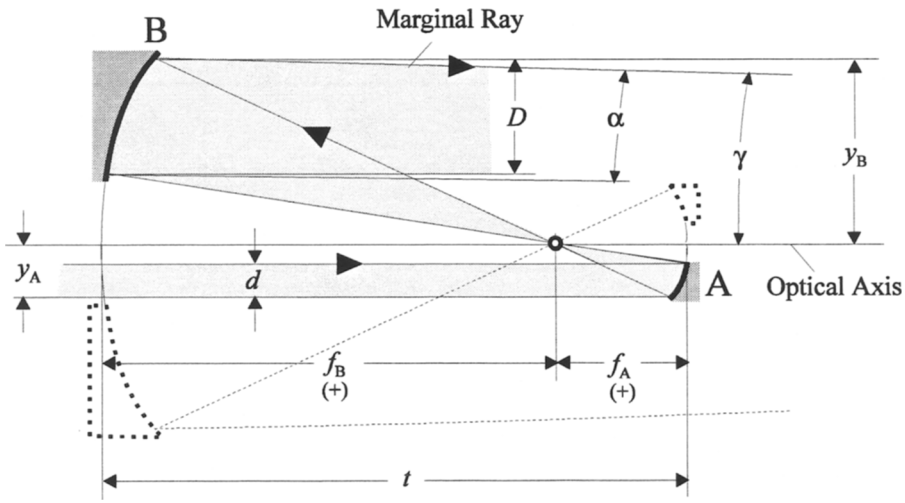


FIG. 3.41 Gregorian beam expander.

The angular convergence for both configurations is

$$\gamma = -\frac{1}{64(f/\#)^3}\left[\frac{m-1}{m}\right], \tag{3.73}$$

where m is the magnification or expansion ratio D/d. It is important to remember that m is positive for the Cassegrain expander and negative for the Gregorian.

The relative aperture $(f/\#) = f_B/(2y_B) = |f_A|/(2y_A)$, the separation of the mirrors $t = f_A + f_B$. For an $f/2$ system with $m = 5$, the convergence angle γ for the Cassegrain expander is $\cong -1.5$ mrad, and for the Gregorian with $m = -5$, $\gamma \cong -2.3$ mrad. Notice that the factor in front of the square bracket of Eq. (3.73) is equal to twice the value of the minimum angular blur spot size of a single spherical mirror due to spherical aberration.

To correct spherical aberration with two reflectors, at least one of them has to be aspheric. For both configurations, the classical Gregorian and the Cassegrain, all mirrors have to be paraboloids. Details will be covered in Chapter 5.

3.8 Diffraction Limit

An ideal optical system would image an object point perfectly as a point. However, due to the wave nature of radiation, diffraction occurs, caused by the limiting edges of the system's aperture stop. The result is that the image of a point is a blur, no matter how well the lens is corrected. This is the diffraction blur or Airy disk, named in honor of Lord George Biddel Airy, a British mathematician (1801–1892). Its cross section and its appearance are shown in Fig. 3.42.

Since this blur size is proportional to the wavelength, as indicated in Eq. (3.74), the diffraction effect can often become the limiting factor for infrared systems. If the aperture of the lens is circular, approximately 84% of the energy from the energy from an imaged point is spread over the central disk surrounded by the first dark ring of the Airy pattern. The diameter of this central disk is

$$B_{diff} = 2.44\lambda\,(f/\#). \tag{3.74}$$

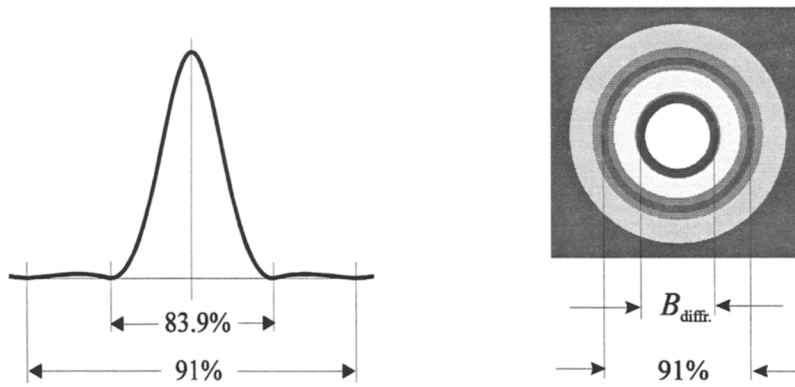


FIG. 3.42 Airy disk, energy distribution and appearance.

For a LWIR system operating at a wavelength of 10 μm and an *f*-number of 2, the diffraction blur with 84% of the energy is 48.8 μm (approximately 0.002 in.) in diameter. The trend to decrease the pixel sizes in focal plane arrays to increase the system's resolution requires awareness of this limit set by nature. Including the energy of a diffraction blur twice in diameter of the central disk (the measure of the second dark ring) increases the total encircled by only 7%. The spread of the energy outside the central disk can cause disturbing crosstalk into adjacent detector elements in an array.

3.9 Resolution of Imaging Systems

Two points are resolved when their images are separated enough from each other so that they can be clearly recognized as two points. A convenient and widely used value for the minimum separation is the radius of the first dark ring of the Airy pattern, as indicated in Fig. 3.43. This is Rayleigh's criterion, named after Lord Rayleigh, a British physicist (1842–1919).

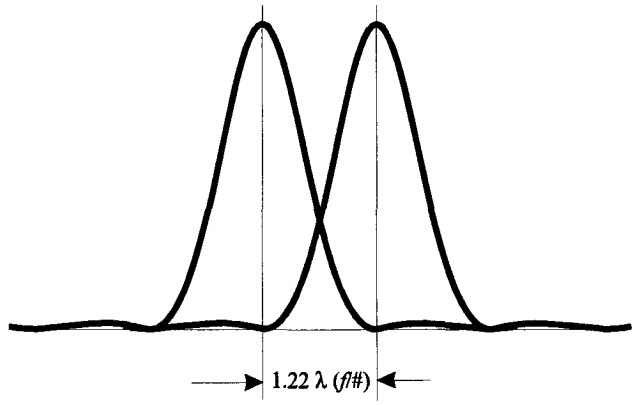


FIG. 3.43 Lord Rayleigh's resolution criterion.

References

1. J. R. Meyer-Arendt, *Introduction to Classical and Modern Optics*, second edition, Practice-Hall, Inc. (1984), page 120.
2. W. T. Welford, *Aberrations of Optical Systems*, Adam Hilger (1989), page 228.
3. K. Mütze, *ABC der Optik*, Verlag Werner Dausien (1961), page 307.
4. W. J. Smith, *Modern Optical Engineering*, McGraw-Hill Inc. (2000), page 335.
5. M. J. Riedl, "The Thin Achromat," Electro-Optical Systems Design, Cahners Publishing Co. (September 1981), pages 49–52.
6. M. J. Riedl, "Astigmatism and the spherical mirror," *Electro-Optical Systems Design,* Milton S. Kiever Publications, Inc. (Sept. 1977), pages 27–31.
7. M. J. Riedl, "The Mangin mirror and its primary aberrations," *Applied Optics,* Vol. 13, No.7, pp. 1690–1694 (1974).
8. W. J. Smith, *Modern Optical Engineering*, McGraw-Hill Inc. (2000), page 328.
9. R. Kingslake, *Lens Design Fundamentals*, Academic Press (1978), page 208.
10. I. A. Artioukov and K. M. Krymski, "Schwarzschild objective for soft x-rays", *Optical Engineering,* (August 2000), Vol.39, No. 8, pages 2163–2170.

CHAPTER 4

Wave Aberrations

4.1 Introduction

So far, it has been assumed that light propagates as rays. This is not the case. Rays only exist in theory and are a convenient basis for geometric optics. In reality, rays are lines drawn in the direction of the propagation of light. They are the normals to the wavefront in the direction of the flow of the radiant energy. This is indicated in Fig. 4.1, where light radiates from a point source P in spherical waves. After exiting an aberration-free optical system, the wavefront converges precisely toward point P'.

If the media through which the light travels are not isotropic, the wavefront is no longer spherical. Advantage is taken of this fact with gradient index material to form the desired wavefront shape.

To explain the concept of wave aberration in this fundamental text, we limit ourselves to the discussion of the spherical aberration. It is sometimes desirable to compare the results obtained with geometrical optics. This will be done for third-order spherical aberration. For an in-depth study, the reader is encouraged to consult references 1 and 2.

4.2 Diverging and Converging Waves

Light emanating from a point source forms a diverging spherical wavefront. After passing through a fitting optical system, it is converted into a converging wavefront, aiming for the axial image point as indicated in Fig. 4.1.

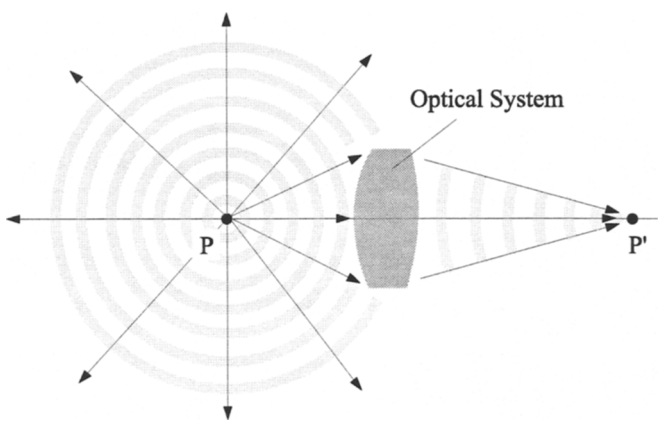


FIG. 4.1 Diverging and converging waves in isotropic media.

4.3 Optical Path Length OPL

The total optical path length through an optical system is the sum of the product Nt along each ray, where N is the refractive index of any portion of the ray path and t is the length of that portion, i.e.,

$$OPL = \sum Nt. \tag{4.1}$$

In Fig. 4.2 two rays are shown; the axial and the marginal ray. The optical paths lengths are, respectively, $OPL_A = t_{A1} + Nt_{A2} + t_{A3}$ and $OPL_M = t_{M1} + Nt_{M2} + t_{M3}$.

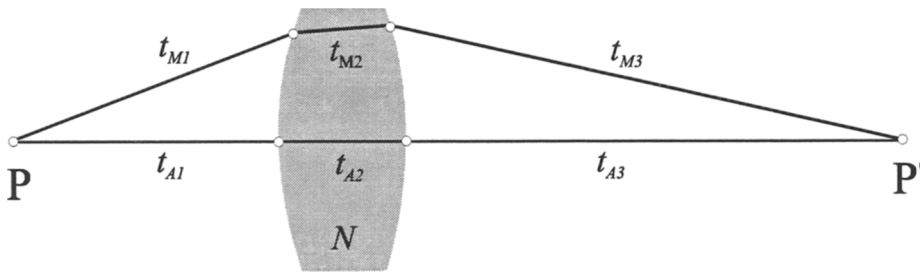


FIG. 4.2 Optical path lengths for axial and marginal rays.

If OPL_M is equal to OPL_A, there is no difference between the two optical path lengths and, therefore, there is no wavefront deformation. The wavefront exiting the lens is spherical and concentric about the image point P'.

4.4 Optical Path Difference OPD (Wavefront Aberration)

Once again referring to Fig. 4.2, if there is a difference between the two optical path lengths, aberration is present. This wavefront aberration is expressed as

$$OPD = \left(\sum Nt\right)_A - \left(\sum Nt\right)_M. \tag{4.2}$$

4.5 Spherical Aberration

If the paraxial focal point of the lens or system is used as a reference point about which a reference sphere is constructed, then the distance between the point where the marginal ray—the normal to the exiting wavefront—and the paraxial focus is the longitudinal spherical aberration. This is shown in Fig. 4.3.

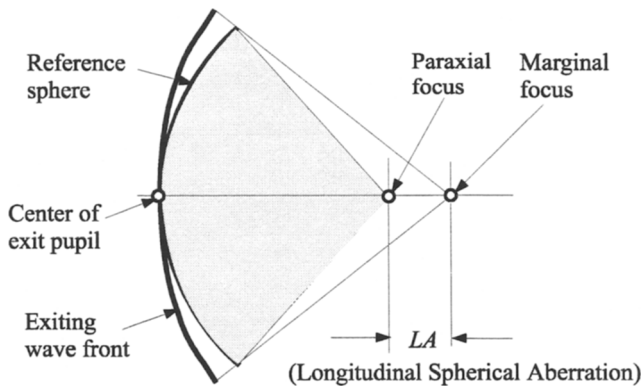


FIG. 4.3 Longitudinal spherical aberration is the distance between the marginal and paraxial foci.

4.5.1 Numerical example

We analyze a doublet made from germanium with index $N = 4$. The chosen focal length $f = 100$ mm, the $f/\# = 2$, and the design wavelength $\lambda = 10$ μm. After optimizing the thin lens configuration, the following prescription is obtained:

Front element :	$R_1 = 232.5$ mm	
		Thickness $t_{1/2} = 4.5$ mm
	$R_2 = 348.5$ mm	
		Spacing $t_{2/3} = 0.5$ mm
Rear element :	$R_3 = 87.0$ mm	
		Thickness $t_{3/4} = 4.5$ mm
	$R_4 = 99.572$ mm	

To calculate the optical path lengths is somewhat cumbersome, but the exercise brings with it a better understanding of the subject under discussion. Figure 4.4 and Table 4.1 provide a summary of this exercise.

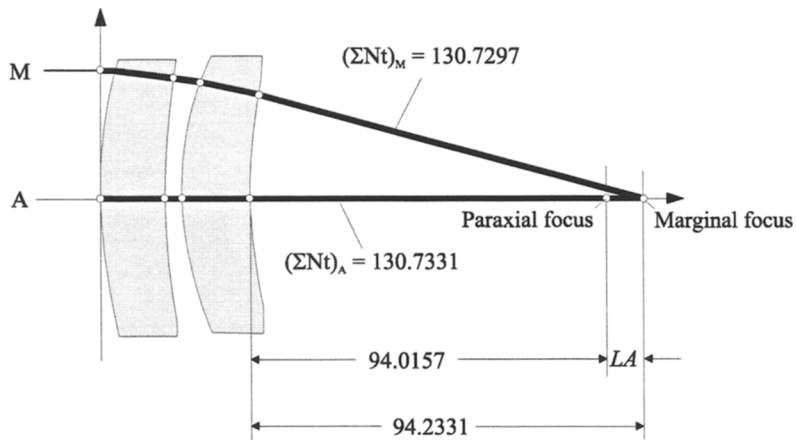


FIG. 4.4 Optical path lengths for rays A and M. (The illustration is not to scale).

Table 4.1 OPLs with respect to marginal focus for doublet shown in Fig. 4.4.

Optical Path Length for Ray M	**Optical Path Length for Ray A**
1.3479940 16.1589969 3.1155744 15.7162411 94.3909159	18.0000000 0.5000000 18.0000000 94.2331377
$(\Sigma Nt)_M = OPL_M = 130.7297223$	**$(\Sigma Nt)_A = OPL_A = 130.7331377$**

The optical path difference OPD = 130.7331−130.7297 = 0.0034 mm, which is equivalent to 0.34 waves for the chosen wavelength of 10 μm. This is shown as plots in Fig. 4.5.

The longitudinal spherical aberration LA = 94.2331−94.0157 = 0.2174 mm. As already stated, this is the distance between the marginal and the paraxial foci.

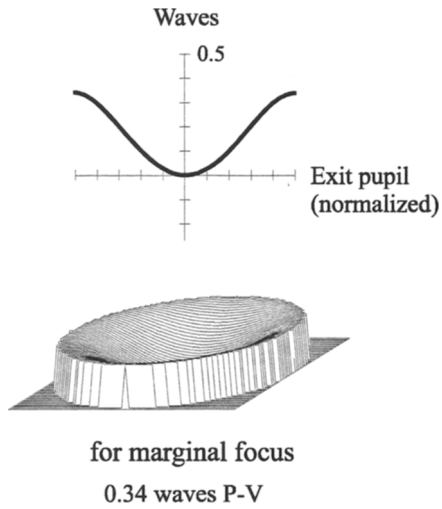


FIG. 4.5 Wave aberration (OPD) with reference to marginal focus.

4.5.2 Best focus position

In Fig. 3.11 we have shown that there is an optimum location for the third-order blur spot due to spherical aberration. That location is based on geometric optics. Based on wave aberration, the optimum position lies halfway between the marginal and paraxial image planes. In other words, by refocusing, the wavefront will be referenced to a sphere whose center lies now halfway between the marginal and paraxial foci. The effect is shown in Fig. 4.6; the plots refer to the same doublet discussed above.

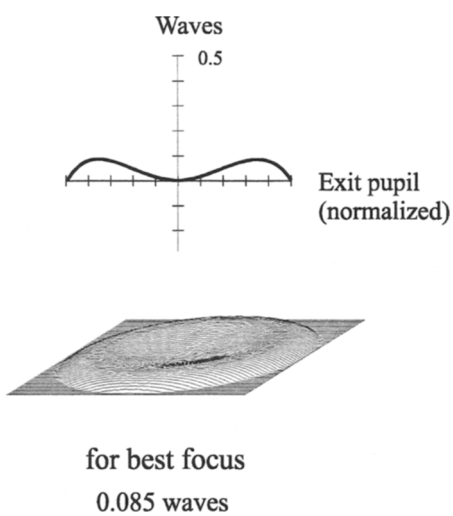


FIG. 4.6 Wave aberration for best focus position.

4.6 Third-Order Spherical Aberration

To simplify the calculations and to show the relation between geometrical and wave aberration and using third-order aberration theory, the following relation exists:

$$OPD_{\text{par}} \cong \frac{LA}{16(f/\#)^2}, \tag{4.3}$$

where OPD_{par} indicates the optical path difference with regard to the paraxial focus.

Applying Eq. (4.3) to our example, with LA = 0.2174 mm and $f/\#$ = 2, the third-order wave aberration is

$$OPD_{par} \cong \frac{0.2174}{16 \cdot 2^2} = 0.0033969\,\text{mm},$$

which is 0.34 waves peak to valley for a wavelength of 10 μm. This agrees well with the result obtained using the more tedious way shown in Eq. (4.2).

For the best image position—which lies, as already stated, half way between the paraxial and marginal focal points—the third-order spherical aberration reduces to

$$OPD_{\text{min}} \cong \frac{LA}{64(f/\#)^2} = \frac{OPD_{par}}{4}. \tag{4.4}$$

This translates for our example into 0.34 / 4 = 0.085 waves and corresponds to the results illustrated in Fig. 4. 5.

4.6 Depth of Focus

Lord Rayleigh found that if the optical path difference of an optical system is not more than one quarter of a wavelength, for all practical purposes, the system can be considered diffraction limited. Therefore, the permissible defocusing range for such a system is

$$\Delta_{\text{focus}} \cong \pm 2\lambda(f/\#)^2. \tag{4.5}$$

This would mean that for our doublet, the allowable defocusing limit would be ± 80 μm.

References

1. W. J. Smith, *Modern Optical Engineering*, McGraw-Hill (2000), pp. 348–352.
2. R. E. Fischer / B. Tadic-Galeb, *Optical System Design,* McGraw-Hill (2000), Chap. 4.

CHAPTER 5

Special Optical Surfaces and Components

5.1 Introduction

An aspheric surface is considered a special optical surface. It is special because it can do more than a spherical surface with regard to aberration correction. But it is also special because it is much more expensive to manufacture, especially if the surface is produced by means other than diamond turning. A diffractive surface with its grating profile engraved into it to act as a thin lens belongs in this category of special surfaces as well. Both will be discussed in detail.

Because windows, filter substrates, and other plane-parallel elements such as beam splitters and Dewar windows are often ignored by the user and not considered to be optical elements, we treat them here as elements with special surfaces. These elements in converging or diverging light contribute to aberrations just as a lens or a curved mirror does. In addition, they shift and displace image locations. Knowing the behaviors of plane-parallel plates allows us to correct for their aberration contributions or take advantage of that knowledge in balancing the system's aberrations.

Another element surprisingly often ignored in the layout stage is the concentric dome. It contributes to the aberrations and has focusing power. The details of these characteristics will be addressed.

Ball lenses and gradient index lenses are frequently used as coupling elements with optical fibers. Their features and performances are included as well, even though their applications are limited to the visible (VIS) and near-infrared (NIR) spectra.

5.2 The Plane-Parallel Plate

By definition, a plane-parallel plate is flat and parallel. If this is not the case, it is a lens with convex or concave spherical or cylindrical surfaces. The surfaces can also be toroidal or general aspheres on either or both sides. The plate can also be flat on both sides, with the surfaces not parallel to each other. In that case the element is a prism.

5.2.1 Displacements

If a plane-parallel plate is inserted into an image-forming optical system, with its surfaces perpendicular to the optical axis, the image plane is shifted toward the right, assuming the light is coming from the left (see Fig. 5.1). The amount of the shift depends

on the index of refraction of the material used. The higher the index, the larger the shift. For the paraxial region, this longitudinal displacement is

$$\Delta_p = \frac{(N-1)}{N} t. \tag{5.1}$$

t is the thickness of the plate with index N. Outside the paraxial region, a more complex relationship applies:

$$\Delta_m = \left[1 - \sqrt{\frac{1-\sin^2 u}{N^2 - \sin^2 u}}\right] t \quad \cong \quad \left[1 - \sqrt{\frac{4(f/\#)^2 - 1}{4N^2 (f/\#)^2 - 1}}\right] t, \tag{5.2}$$

where $\tan u = \dfrac{1}{2(f/\#)} \approx \sin u$. While this introduces an error of about 12% for an $f/1$ cone, the error is already down to 3% for $f/2$ and less than 1% for $f/4$.

The difference between Δ_m and Δ_p is the longitudinal spherical aberration contributed by the plane-parallel plate. It is a positive value, which is referred to as over-corrected spherical aberration. We remember that a single element has under-corrected spherical aberration. In other words, the combination improves the situation slightly.

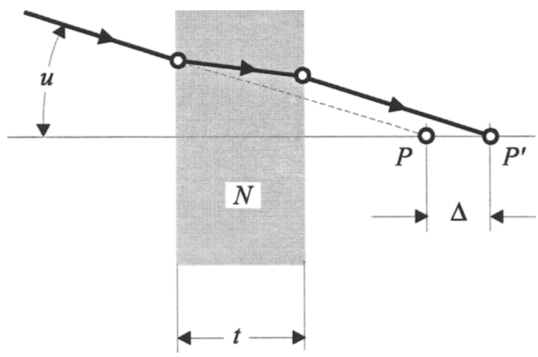


FIG. 5.1 Longitudinal shift of image by plane-parallel plate.

The second type of displacement occurs when the plate is tilted relative to the optical axis. The tilt causes the image to be shifted laterally, as seen in Fig. 5.2.

$$\Gamma_m = \left[1 - \sqrt{\frac{1-\sin^2 u_p}{N^2 - \sin^2 u_p}}\right] t \sin u_p \,. \tag{5.3}$$

Often, a beamsplitter is tilted 45° relative to the optical axis. The offset then becomes

$$\Gamma_{45} = \sqrt{0.5}\left[1-\left(2N^2-1\right)^{-1/2}\right]t. \tag{5.4}$$

This lateral displacement is about one third of the thickness for glass with $N = 1.5$ and more than one half of t for germanium with $N = 4$.

For small tilt angles u_p,

$$\Gamma_s = \frac{(N-1)}{N} t u_p = \Delta_p u_p \ . \tag{5.5}$$

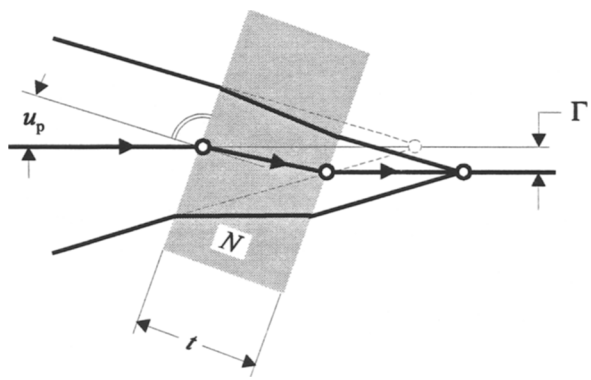


FIG. 5.2 Lateral displacement of image by a tilted plane-parallel plate.

5.2.2 Optical micrometer

Frequently, a plane-parallel plate is employed to measure or set lateral displacements. For a chosen thickness and material, the displacement is only dependent on the sine of the tilt angle. In alignment telescopes, for example, two such plates are used orthogonal to each other to measure horizontal and vertical shifts (see Fig. 5.3). Sometimes, a detector assembly is too heavy and bulky with its cooling system and preamplifier package to be aligned, and the two-plate tilt method can be used to align the optical axis of the system to the detector.

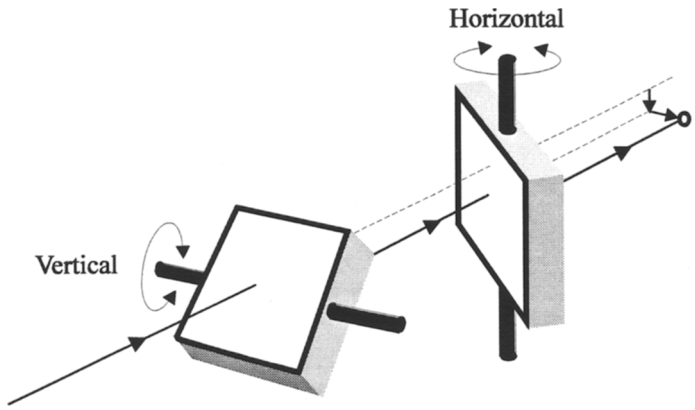


FIG. 5.3 Principle of an optical micrometer.

For a 20-mm-thick plate made from BK7 with index 1.517, the following relationship exists:

Plate Rotation Angle u_p (deg)	1.681	3.358	5.029	6.690	8.339
Lateral Displacement Γ (mm)	0.2	0.4	0.6	0.8	1.0

5.2.3 Aberration contributions

The linear aberration blurs introduced by a plane-parallel plate, based on third-order theory, are

spherical aberration $$B_{\text{spher}} = \frac{(N^2-1)t}{32N^3(f/\#)^3}, \tag{5.6}$$

coma $$B_{\text{coma}} = \frac{(N^2-1)tu_p}{8N^3(f/\#)^2}, \tag{5.7}$$

astigmatism $$B_{\text{astig}} = \frac{(N^2-1)tu_p^2}{2N^3(f/\#)}, \tag{5.8}$$

axial chromatic aberration $$B_{\text{chrom}} = \frac{(N-1)t}{2N^2\nu(f/\#)}, \tag{5.9}$$

and lateral color $$B_{\text{lat.color}} = \frac{(N-1)tu_p}{N^2\nu}. \tag{5.10}$$

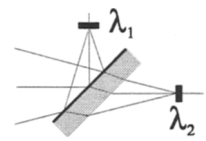
As indicated in Figs. 5.1 and 5.2, u_p is the tilt angle of the plate, and N is the refractive index of the plate material. The relative aperture $(f/\#) = 1/(2u)$.

5.2.4 Application remarks

Let us look at a beamsplitter application in the 3- to 5-μm region. The beamsplitter is inserted into an $f/1.5$ cone, the tilt angle is 45°, and the thickness of the beam-splitter is 1 mm. We will compare two different materials, silicon and sapphire.

The index of refraction at 4 μm is 3.425 for silicon and 1.675 for sapphire. Their Abbe numbers are 235 and 7.6 respectively. This large difference between the Abbe numbers in this region should be noted.

Material	B_{spher}	B_{coma}	B_{astig}	B_{chrom}	$B_{lat.color}$	B_{total}
Silicon	2.5	11.6	54.9	0.3	0.7	70.0 μm
Sapphire	3.6	16.8	78.9	10.5	24.9	134.7 μm



The spot diagrams shown in Figs. 5.4, and 5.5, obtained with geometrical ray traces, demonstrate at once an excellent agreement with the third-order calculations from above.

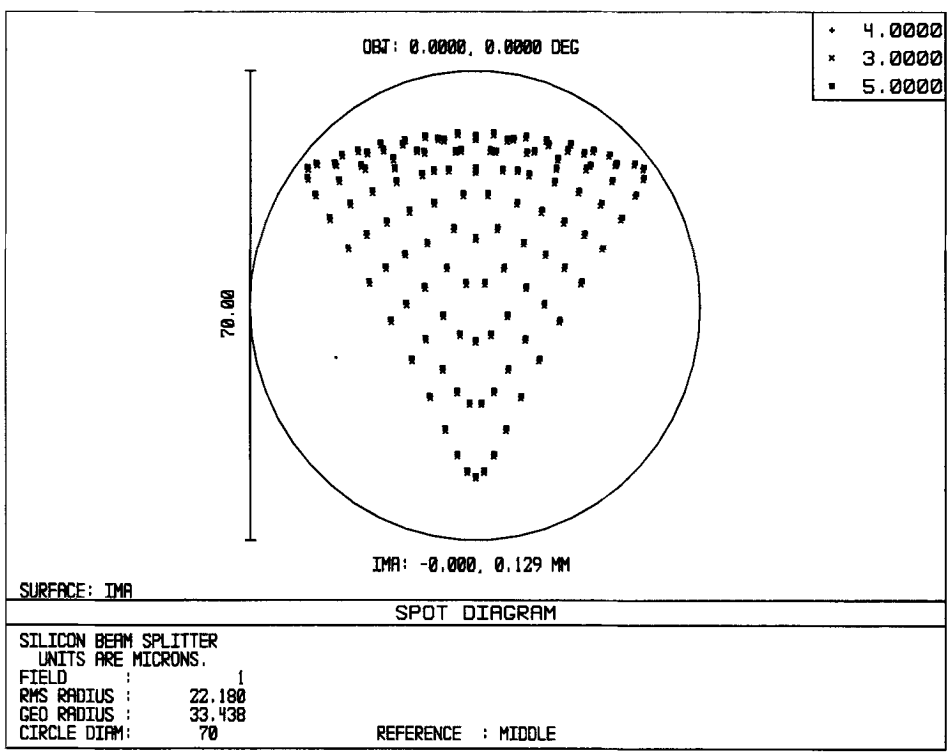


FIG. 5.4 The spot diagram resulting from the silicon beamsplitter.

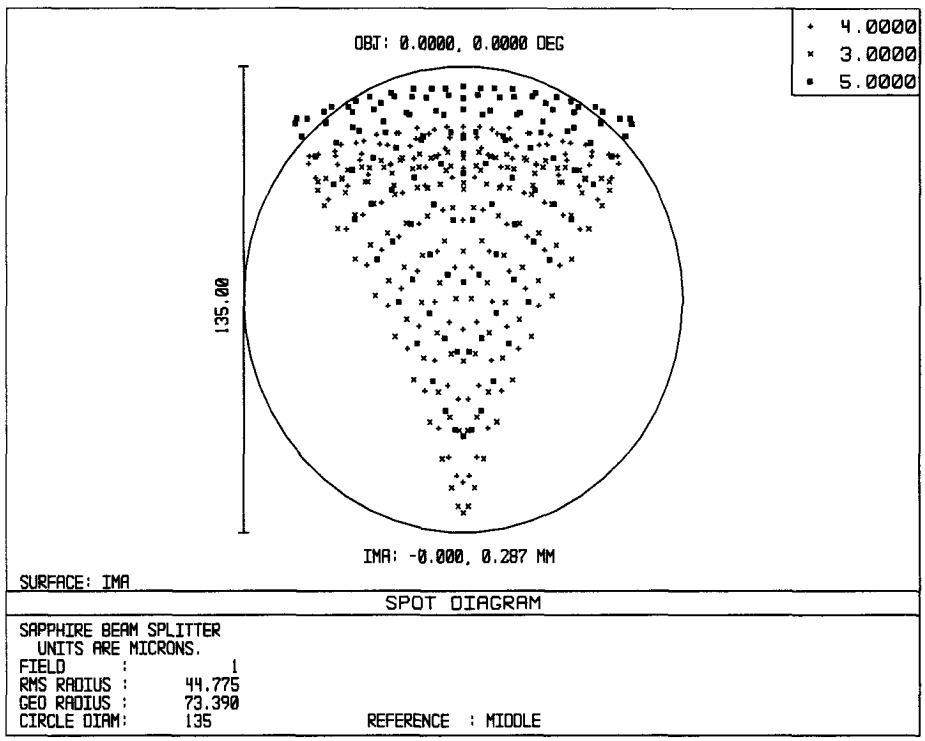


FIG. 5.5 The spot diagram resulting from the sapphire beamsplitter.

This little exercise indicates that the contribution of a relatively thin plane-parallel plate cannot be ignored. Moreover, the comparison of the materials points out the danger in quickly substituting one material for another without considering the consequences.

As mentioned earlier, filters with various thicknesses are frequently inserted into a well-balanced optical system without consideration of the effects discussed above. These effects differ, depending on where those filters are placed within the optical train. A few examples are shown in Fig. 5.6. In the upper layout, the case of finite conjugates is presented with S and S' indicating object and image distances. Notice the apparent shift Δ of the object when the filter (plane-parallel plate) is inserted on the object side, as is indicated in the center arrangement. In this case, the lens "sees" the object closer than without the filter and shifts the image to the right by δ, which also changes the magnification of the system. In the bottom sketch, the filter is placed to the right of the lens, which introduces the shift Δ. The magnification is not affected. It should be clear that a plane-parallel plate inserted in a collimated bundle of light does not contribute any aberrations. However, if the plate is tilted, there is, of course, a lateral displacement of the beam but not the image.

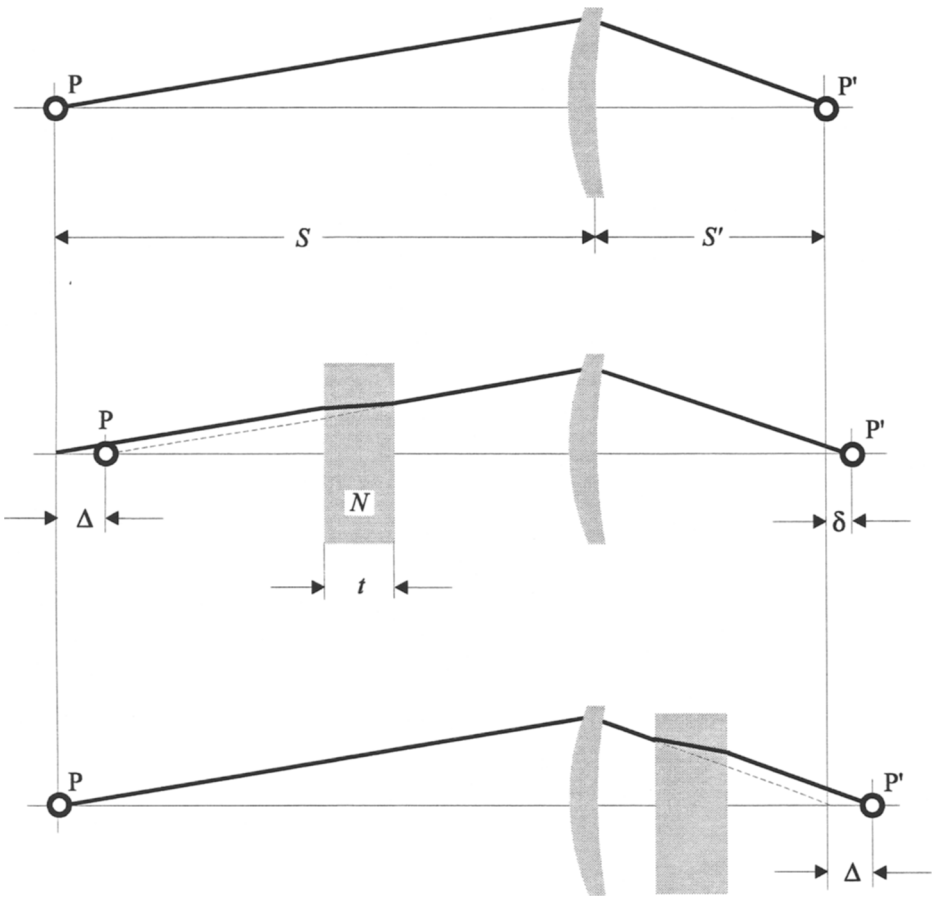


FIG. 5.6 Effects of inserting a plane-parallel plate into a converging or diverging beam.

As before, $\Delta = \dfrac{(N-1)}{N}\, t$.

$$\delta \cong m^2 \Delta = \left[\frac{S'}{S}\right]^2 \Delta = \left[\frac{f}{S+f}\right]^2 \Delta \ . \tag{5.11}$$

For a system with a focal length of $f = 100$ mm, $S = -400$ mm, $t = 3$ mm, $N = 4$, and $\Delta = 2.25$ mm, the shift $\delta \cong 0.25$ mm.

A better arrangement for the application illustrated in Fig. 5.7 is shown in Fig. 5.8. The objective is split into two elements, both chosen for infinite conjugate. Several advantages can be gained from such an arrangement:

1. Each lens can be used for focusing without changing the system magnification.
2. The distance from the object to the image $\overline{PP'}$, which is called the track length, is flexible within reason, depending on the object size.

3. The parallel beam between the two elements allow for better filter performance.
4. There is no image shift with change of index variation of the filter material.
5 Variation of filter thickness does not have any influence on image position or magnification. This is a very convenient advantage when filters with different thicknesses are inserted during measurements.
6. Aberrations are reduced by splitting the lens into two elements. Each can be shaped for optimum performance.
7. The filter or beamsplitter can be tilted without introducing astigmatism.

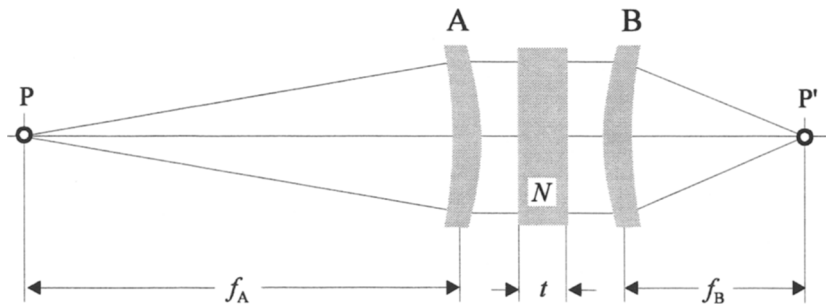


FIG. 5.7 A better arrangement than shown in Fig. 5.6. Magnification $m = -f_B/f_A$.

A similar arrangement is presented in Fig. 5.8 for a reflective system. Here, the filter is not only in parallel light, its size is the smallest possible.

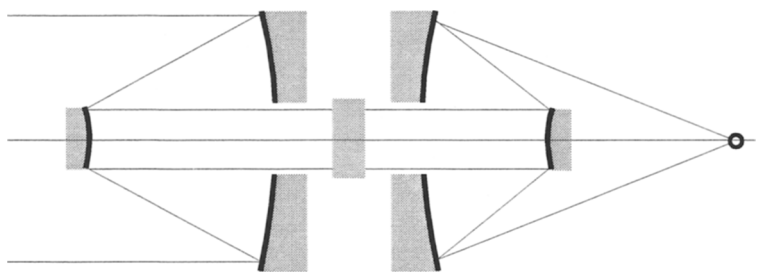


FIG. 5.8 Reflective system with same principle applied as in Fig. 5.7.

5.2.5 The wedge (thin prism)

Two flat surfaces, slightly out of parallel to each other, form a wedge. Applying Snell's law in determining the direction a ray will take after passing through the wedge, and assuming that the wedge angle is small, we find the deviation from the entering ray direction to be

$$\delta = \alpha (N - 1). \tag{5.12}$$

The dispersion is

$$d\delta = \frac{\delta}{v}, \tag{5.13}$$

where $V = \dfrac{N_M - 1}{N_S - N_L}$. All other symbols are shown in Fig. 5.9.

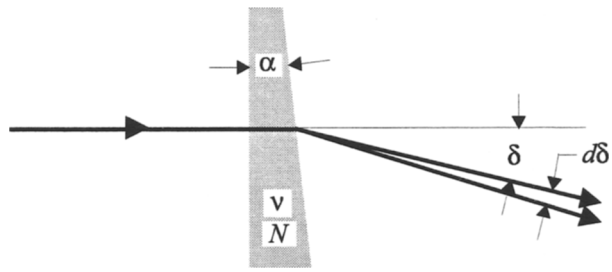


FIG. 5.9 Deviation and dispersion of a thin prism (wedge).

A useful combination of lens plus wedge, as shown in Fig. 5.10, has been applied by the author in a simple two-color discriminator. Two wedges installed behind a common singlet, shaped for minimum spherical aberration, provide a cost-effective method to split the aperture and create two subsystems, each working in a specific spectral band. The wedges serve as filter substrates and redirect the focused energies for clear separation, yet are close enough so that the two detector elements of different chemistry can be installed into a single hermetically sealed standard detector housing.

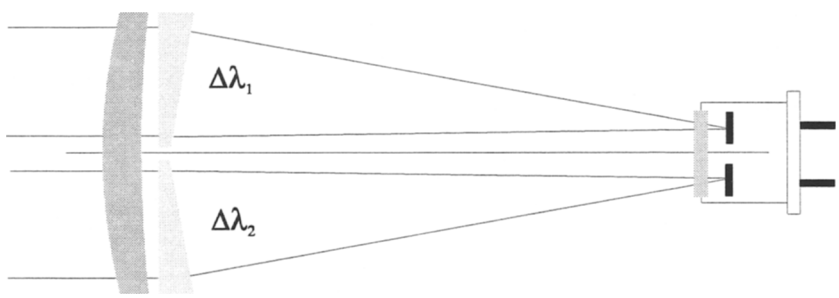


FIG. 5.10 A two-color discriminator optical system.

5.3 Domes

Domes are mostly monocentric shells, which means that the second, or inside, radius, is shorter by the amount of the element's thickness than the first, or outside radius (see Fig. 5.11). Such an element has a focal length of

$$f_{\substack{\text{conc.}\\ \text{dome}}} = -\frac{NR_1(R_1 - t)}{(N-1)t}. \tag{5.14}$$

It is a negative element and contributes an overcorrected spherical aberration. Of course, one can take advantage of this fact in balancing the total system aberrations.[1]

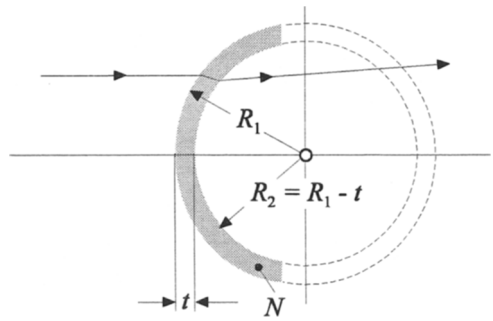


FIG. 5.11 Concentric dome is a negative element.

A dome with zero power can be derived by setting Eq. (3.2) to zero, i.e.,

$$\frac{1}{f} = (N-1)\left[\frac{1}{R_1} - \frac{1}{R_2} + \frac{(N-1)t}{N R_1 R_2}\right] = 0. \tag{5.15}$$

Solving for R_2 yields that R_2 is equal to R_1 less the axial optical displacement caused by the thickness of the dome, or

$$R_2 = R_1 - \frac{(N-1)}{N} t. \tag{5.16}$$

If this concept of a zero power lens is extended to a very thick element, a beam expander evolves (see Fig. 5.12). The magnification of a beam expander is the ratio between the exiting and entering beam diameters. With this definition, the thickness of a solid beam expander can be stated by

$$t = \frac{N}{(N-1)}(1-m)R_1. \tag{5.17}$$

The second radius is simply

$$R_2 = mR_1, \tag{5.18}$$

i.e., the magnification is the ratio of the two radii, independent of the lens material.

For example, a solid 4× germanium beam expander (N = 4) has a first radius of −10 mm. Therefore, its thickness according to Eq. (5.17) is 40 mm, and its second surface radius is −40 mm [Eq. (5.16) or (5.18)]. The spherical surfaces introduce spherical aberrations. The correction for these will be discussed in Section 5.6.

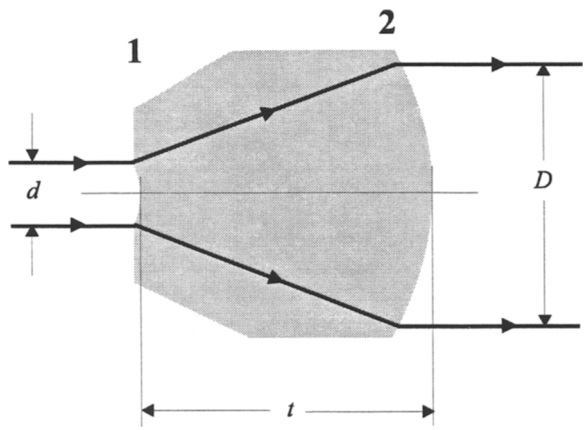


FIG. 5.12 Solid beam expander with $m = D/d$.

5.4 The Ball Lens

A sphere or ball performs surprisingly well as a lens. At closer scrutiny, one finds that such an element can be broken down into two plano-convex lenses, separated by a plane-parallel plate. The positive lenses have under-corrected spherical aberration and the plane-parallel plate is over-corrected. Therefore, there is a compensating effect. A ball lens finds its main application as a coupling element for optical fibers. It deserves an analysis, especially since there are certain limitations that need to be understood.

Figure 5.13 indicates that the focal length is measured from the center of the sphere, where the extensions of the entering and exiting rays meet. This means that the back focal distance *bfl* is merely the difference between the focal length and the radius R of the sphere. One can see immediately that the focus falls on the rear surface of the sphere if $f = R$.

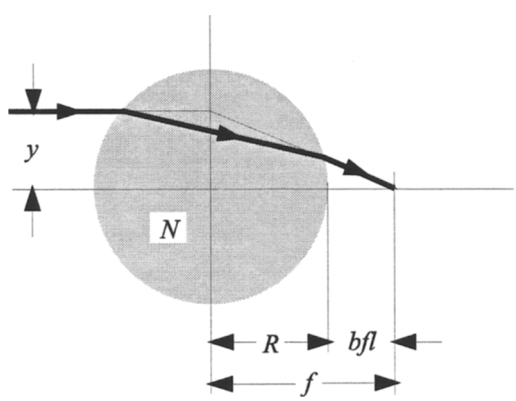


FIG. 5.13 Ray trace through a ball lens.

The focal length of the ball lens is

$$f = \frac{RN}{2(N-1)}. \tag{5.19}$$

Its back focal length or back focal distance is, as already indicated,

$$bfl = f - R = \frac{(2-N)R}{2(N-1)}. \tag{5.20}$$

One can see that $bfl = 0$ if $N = 2$. That means a ball lens is not suitable for most IR materials, such as silicon, germanium, zinc sulfide, or zinc selenide, whose index of refraction are larger than 2. Sapphire with $N = 1.6753$ at 4 μm is an exception.

5.4.1 Spherical aberration

The transverse third-order spherical aberration is expressed by

$$TSC_{ball} = \frac{[N(N-3)+1]y^3}{2N^2R^2}, \tag{5.21}$$

and since the angular blur spot size, refocused for the best position is $\beta_{spher} = \frac{TSC}{2f}$,

$$\beta_{spher_{ball}} = \frac{N(N-3)+1}{128(N-1)^2(f/\#)^3}. \tag{5.22}$$

This indicates that for a sapphire ball, the angular spherical aberration, measured in radians, is $\beta_{sapp.} \cong 0.034/(f/\#)^3$, which is about 18% better than a lens, made from the same material and shaped for minimum spherical aberration.

5.4.2 An aspherized ball lens

To eliminate spherical aberration, a thin epoxy shell can be attached. This is usually done by a method called *optical replication.* It is a transfer process from a highly polished mold that has the opposite shape of the desired asphere. Figure 5.14 shows such an aspherized sphere. The correcting shell has been placed at the front surface. Because epoxy does not transmit in the MWIR and LWIR regions, this method of aspherizing is limited to the shorter wavelengths, including the NIR. Aspherizing by diamond turning will be discussed in Chapter10.

Before the ball is aspherized, it should be ground into a cylinder, as indicated in Fig. (5.14). This is necessary to establish a centerline for the rotationally symmetrical asphere.

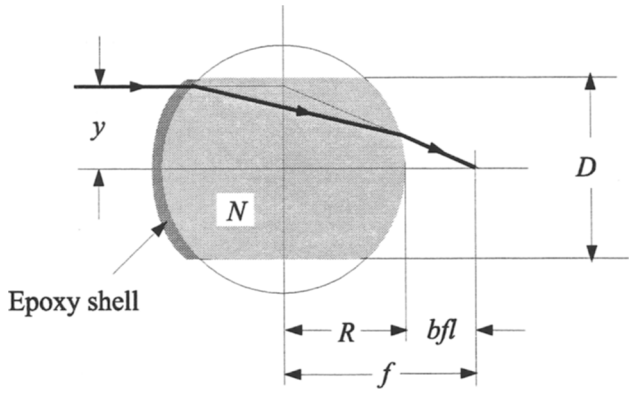


FIG 5.14 Ball lens with aspheric epoxy shell added to correct spherical aberration.

5.5 Gradient Index Lens

Instead of a ball lens, a gradient index lens is frequently used as a fiber coupling element. Such a lens has a gradient profile in which the refractive index varies in the direction perpendicular to the optical axis, as expressed in Eq. (5.23).

$$N = N_0\left[1-(k/2)r^2\right], \tag{5.23}$$

where N_0 is the base index (at the center of the lens), k is called the *Gradient Constant*, and r is the variable radius (mm). Figure 5.15 shows a gradient index rod (GRIN rod) with the length of one full sinusoidal path, or one "pitch."

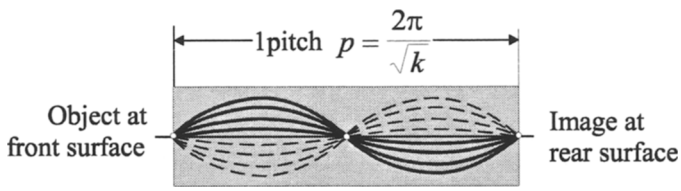


FIG. 5.15 A GRIN rod, one "pitch" long.

Shortening the rod to a length of 1/4 pitch, as shown in Fig. 5.16, forms a lens with a focal length of

$$f = \frac{1}{N_0\sqrt{k}} \ . \tag{5.24}$$

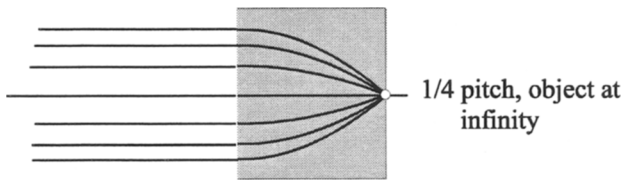


FIG. 5.16 The gradient index lens.

Lenses of that kind are commercially available under the trade name SELFOC, a derivative from "self-focusing."

If the length of such a lens is identified as t, then the focal length is expressed by [2]

$$f = \frac{1}{N_0 \sqrt{k} \sin\left(t\sqrt{k}\right)}. \tag{5.25}$$

The back focal length is

$$bfl = f \cos\left(t\sqrt{k}\right). \tag{5.26}$$

Figure 5.17 shows the details of a GRIN lens with a central index of refraction of $N_0 = 1.5834$ and a gradient constant of $k = 0.1067$. With a length $t = 4$ mm, we obtain by using Eqs. (5.25) and (5.26) a focal length of $f = 2$ mm, and a back focal length of $bfl \approx 0.52$ mm.

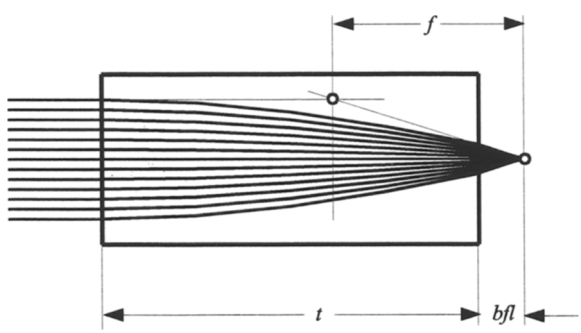


FIG. 5.17 Details of a GRIN rod lens for an object located at infinity. Such an element is also known as a *Wood lens*.

Figure 5.18 shows the same lens as Fig. 5.17, but now the object is located at a finite distance. To better understand the function of a GRIN lens, an "equivalent" thin lens has been superimposed for reference. This clarifies the object and image distances and, with the fields added, the magnification $m = s'/s = h'/h$.

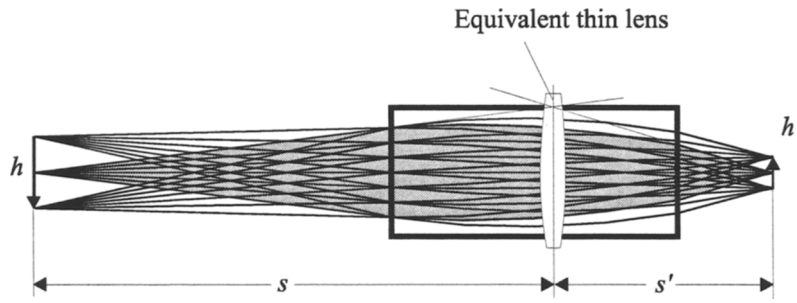


FIG.5.18 A GRIN lens imaging an object located at a finite distance.

5.6 Conic Sections and General Aspheres

Conic sections are called out in Fig. 5.19. They are of special interest to the optical designer. Their distinct characteristics allow the elimination of spherical aberration for refractive as well as reflective surfaces. That is why they are widespread for IR applications. Furthermore, although it is still costly to produce aspheres in glass, it is quite economical to machine them for the MWIR and LWIR, even in very small quantities, by single point diamond turning. Materials such as germanium, zinc sulfide, zinc selenide, and Amtir lend themselves nicely to this process. Metal mirrors, mostly nickel-plated aluminum structures, fall into the same category.

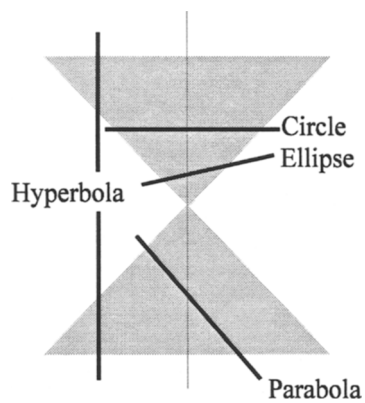


FIG. 5.19 Conic sections.

For the sake of completeness we include the circle as a conic section in our general discussion. The others are the ellipse, the parabola, and the hyperbola. Our interest lies, of course, in surfaces that are generated by rotating these cross sections about an axis of symmetry.

5.6.1 Mathematical expressions

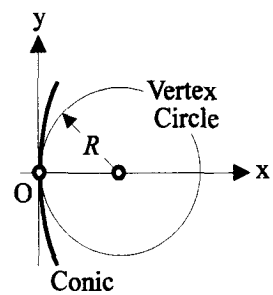
The general equation for a conic section going through the origin of the Cartesian coordinate system is given by

$$y^2 - 2Rx + px^2 = 0, \tag{5.27}$$

with p being the conic constant, sometimes referred to as the aparabolic constant. The relation to the aspheric constant K, also often called the conic constant, is $K = p-1$. R is the radius of the vertex circle.

p determines the shape of the conic section.

For	$p > 1$	Ellipse (oblate),
	$p = 1$	Circle,
	$1 > p > 0$	Ellipse (prolate),
	$p = 0$	Parabola,
	$p < 0$	Hyperbola.



Explicit expressions for a conic section are

$$x = \frac{R - \sqrt{R^2 - py^2}}{p}, \tag{5.28}$$

and

$$x = \frac{cy^2}{1 + \sqrt{1 - pc^2y^2}}, \tag{5.29}$$

with $c = 1/R$. Equation (5.29) is often preferred because it includes the special case of a flat surface, for which $c = 0$.

Frequently, especially for manufacturing purposes, it is desirable to know the radius of a sphere that contacts the conic section at the center and at the edges as indicated in Fig. 5.20. This sphere is called the *best-fit sphere* (BFS). Its radius can be calculated from

$$\rho_{BFS} = R + 0.5(1-p)x_{max} = R + \frac{(1-p)}{4p}\left[2R - \sqrt{4R^2 - pD^2}\right]. \tag{5.30}$$

Also of interest is the slope angle at any given point on the conic surface (see Fig. 5.21). This is simply the first derivative of the curve's equation, which is

$$\frac{dy}{dx} = \tan\alpha = \frac{R - px}{y}, \tag{5.31}$$

where R is the vertex radius of the conic surface.

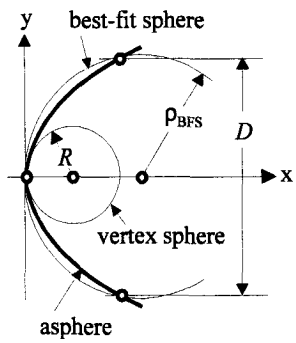


FIG. 5.20 Vertex sphere and best-fit sphere.

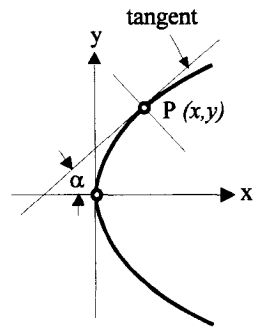


FIG. 5.21 Slope angle α.

5.6.2 Reflectors with conic section surfaces

Applied as reflectors, the family of conic sections can be described as an evolution from the circle. An object placed at the center of the circle will be imaged at the same spot. This is indicated in Fig. 5.22. F_1 is imaged at F_2; as F_1 is displaced some finite distance to the left, the conic section becomes an ellipse. Next, we let F_1 move out to infinity, and the shape of the section turns into a parabola. Finally, one can think of F_1 reappearing from infinity on the right sight of the illustration, where it becomes a virtual object imaged by the hyperbola at F_2. The diagram indicates that the second branch of the hyperbola is also a valid optical surface and can be used to form an aberration-free on-axis image.

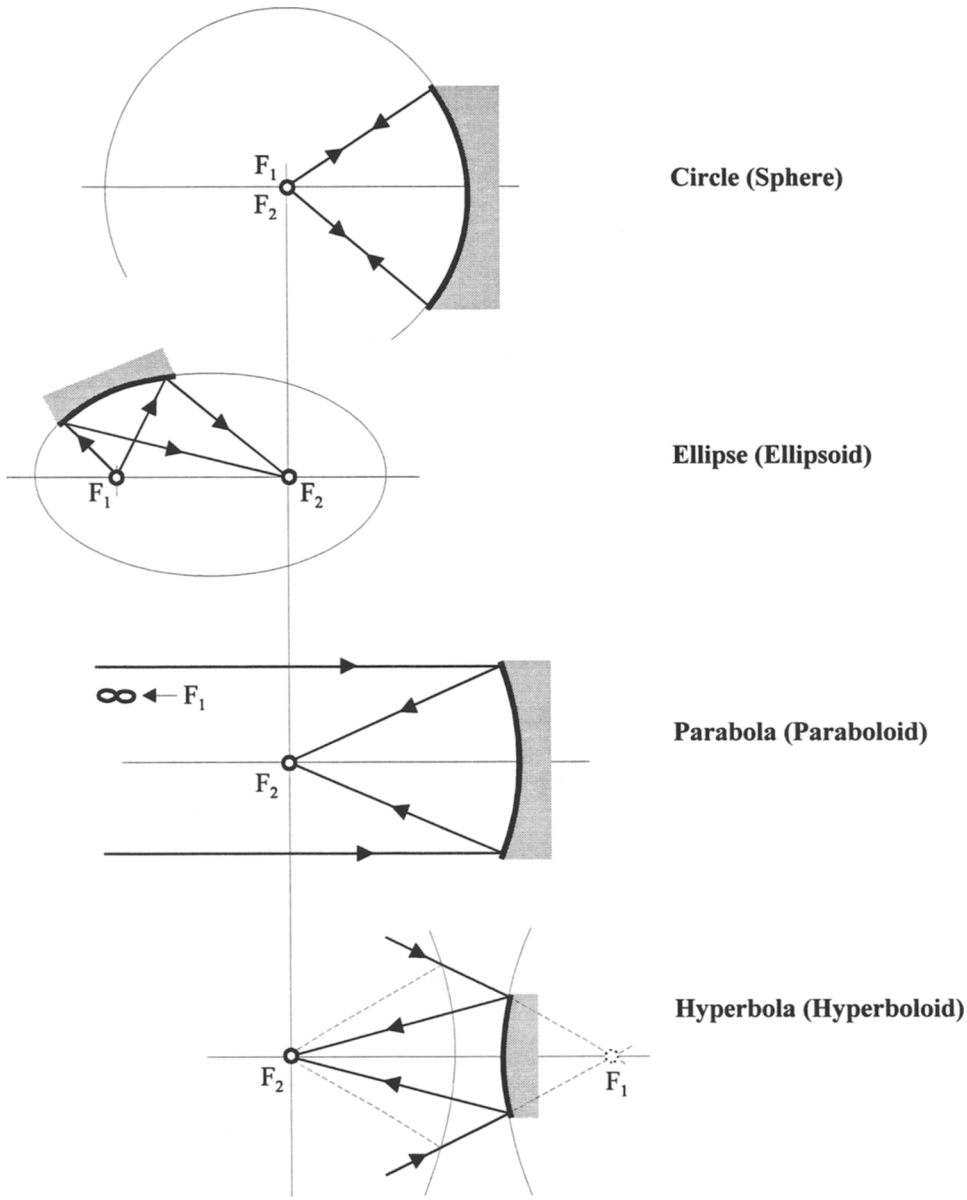


FIG. 5.22 Conic section reflectors.

5.6.3 Lenses with conic section surfaces[3]

Demanding equal optical paths from a plane reference surface to the focal point of a lens, for all rays of a light bundle, yields a lens shape that eliminates spherical aberration. One is shown in Fig. 5.23, where the front surface is an ellipsoid and the second surface is concentric to the focal point of the lens. The spherical aberration is corrected by the front surface alone, and the lens thickness is immaterial.

With reference to Fig. 5.23, the simple relationships for the surface radius, the conic constant, and the ellipse half-axis dimensions are

$$R_1 = \frac{(N-1)}{N} B, \tag{5.32}$$

and

$$p = \frac{(N^2-1)}{N^2}. \tag{5.33}$$

The large half axis is

$$a = \frac{N}{(N+1)} B, \tag{5.34}$$

and the small half axis is

$$b = \sqrt{\frac{(N-1)}{(N+1)}}\, B. \tag{5.35}$$

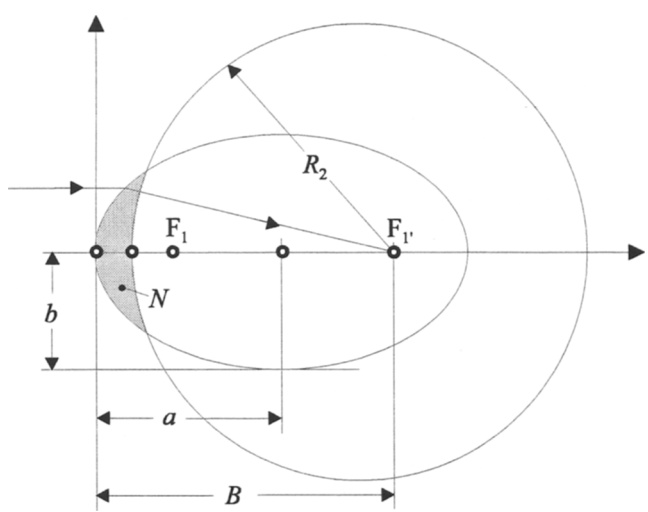


FIG 5.23 Singlet with elliptical front surface and spherical rear surface.

Another singlet, corrected with one conic surface, is a plano-convex type with the curved surface being a hyperboloid. This configuration is illustrated in Fig. 5.24. Again, the parameters for the hyperbola are derived by maintaining equal path lengths for all rays across the lens diameter.

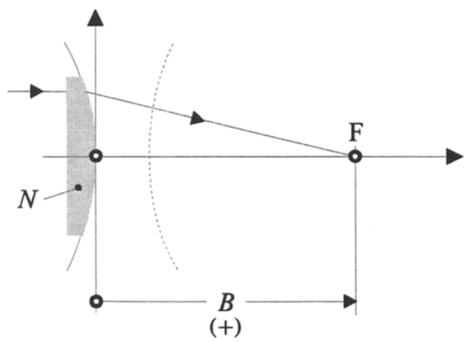


FIG. 5.24 Positive singlet with plano front surface and hyperbolic rear surface.

The relations for the second surface radius and the conic constant are

$$R_2 = (1 - N)B, \tag{5.36}$$

and

$$p = 1 - N^2. \tag{5.37}$$

These equations are also valid for a negative lens as shown in Fig. 5.25. One just has to observe that B has a negative value in this case.

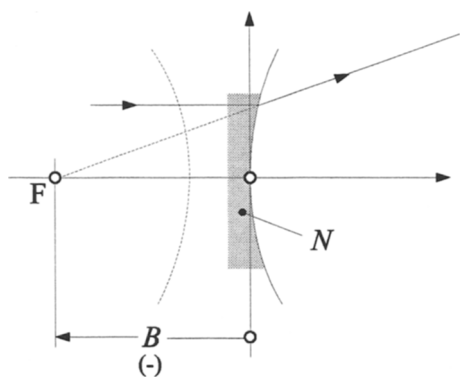


FIG. 5.25 Negative singlet with plano front surface and hyperbolic rear surface.

It can easily be seen that combining two such lenses into one and selecting the appropriate hyperbolic shapes will yield a singlet corrected for spherical aberration for the case of finite conjugates, as indicated in Fig. 5.26.

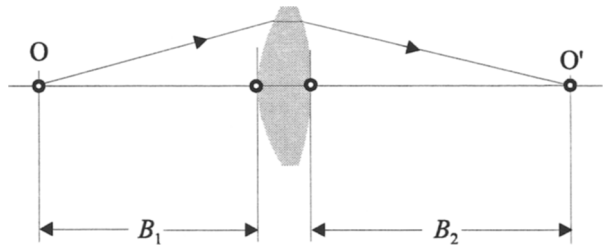


FIG. 5.26 Singlet, with two hyperbolic surfaces, corrected for spherical aberration, applied to finite conjugates.

5.6.4 Common two-mirror configurations using conic section surfaces

There are a number of two-mirror systems that have become standards. These are of special interest for IR applications. In principle, they are alike because they have a concave primary mirror and a convex secondary mirror. The Gregorian arrangement is different. As was pointed out in Chapter 3, its secondary mirror is also concave.

1. The classical *Cassegrain* arrangement consists of a parabolic primary and a hyperbolic secondary mirror. This configuration is corrected for spherical aberration and is, of course, without chromatic aberration.

2. Because a convex aspheric mirror is difficult to measure, a system was developed that uses a spherical secondary mirror. The spherical aberration correction is achieved by shaping the primary mirror into a prolate ellipsoid. This configuration is named after its inventors, *Dall-Kirkham.*

3. A system that is corrected for third-order spherical aberration and coma is achieved by using two hyperbolas. It is called a *Ritchy-Chretien* system.

The equations for the basic radii of the mirrors have been presented in Chapter 3, as Eqs. (3.57) and (3.58). Expressions for the conic constants and the uncorrected aberrations of the named configurations can be found in Ref. 4.

5.6.5 General aspheres (surfaces of rotation)

In modifying a basic conic surface, higher-order deformation coefficients are added to the equation. This provides an extra tool for the correction of some aberrations.

The general rotationally symmetric asphere is presented by

$$x = \underbrace{\frac{cy^2}{1+\sqrt{1-pc^2y^2}}}_{\text{Conic}} + \underbrace{A_4y^4 + A_6y^6 + \ldots + A_jy^j}_{\text{Deformation}} \quad . \tag{5.38}$$

There are two special cases: first, if $p = 0$, the base surface is a sphere; second, if the curvature of the surface is a flat ($c = 0$) with only the deformation coefficients applied.

5.6.6 Two conic section mirrors with a general aspheric corrector

This arrangement allows the correction of third-order spherical aberration, coma, and astigmatism. It is depicted in Fig. 5.27. Because the corrector is a refractive element, some chromatic aberration will be introduced. However, since the corrector is inserted close to the focal plane and its thickness is small, the effect is minimal. The given equations have been developed by using the standard third-order surface contribution expressions.[4]

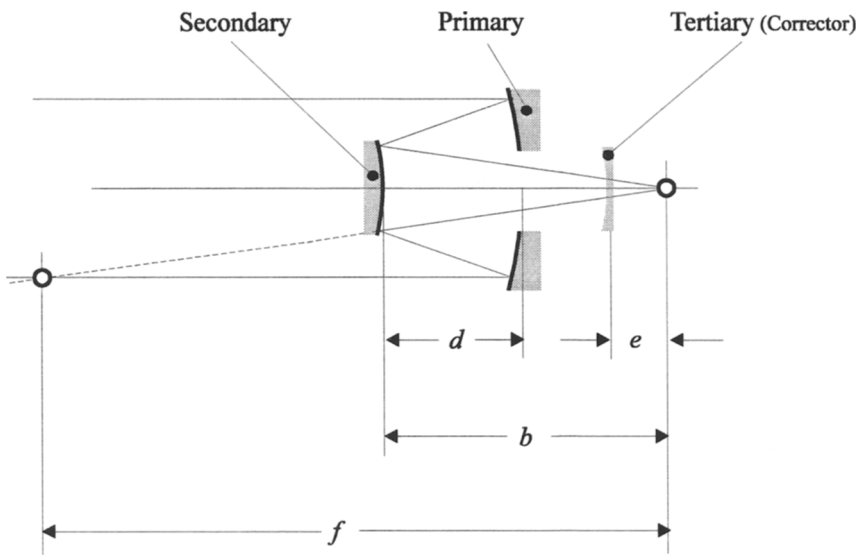


FIG. 5.27 Two conic section mirrors with aspheric corrector.

The conic constants for the primary and secondary mirrors can be calculated from

$$p_1 = \frac{R_1^3}{8d^3f^3}\left[-(b-f)^3 + \frac{2bd^2f(b+e)}{(de-ef+bf)}\right]+1 \tag{5.39}$$

for the primary mirror, and

$$p_2 = \frac{R_2^3}{8d^3b^3}\left[2f(b-f)^2 + (f-d-b)(f+d-b)(d-f-b) + \frac{2d^2e(d-2f)}{(e-b)}\right]+1 \tag{5.40}$$

for the secondary mirror. For the tertiary element, the corrector, the fourth-order deformation coefficient is

$$A_{4_3} = \frac{b(d-2f)}{16e^2(1-N)(e-b)(de-ef+bf)} . \tag{5.41}$$

N is the refractive index of the corrector material.

As elements are added to a system, developing closed solutions for the prediction of the aberrations becomes more complicated. That is where the computer comes into the picture. One begins with a promising fundamental arrangement and uses optimization routines to find a useful solution. The trick is to know what a promising arrangement is, for which there is no substitute for experience.

5.6.7 Three-mirror configurations

It has been already mentioned that due to their wavelengths insensitivity, mirror configurations are frequently favorite candidates for infrared systems. Therefore, some limited remarks about three-mirror systems will be included in this fundamental text.

As with refractive systems, increasing the number of elements in an all-reflective system provides additional opportunities to eliminate or correct more aberrations. However, the simplicity which still exists for most two-mirror systems is quickly lost and with that, it becomes increasingly more difficult to find an optimum configuration to satisfy given requirements.

Figure 5.28 shows an initial three-mirror layout. It is important to notice that all three mirrors are centered about a common axis of rotation, which is also the optical axis. The three mirrors are parabolas and, therefore, spherical aberration is completely corrected.

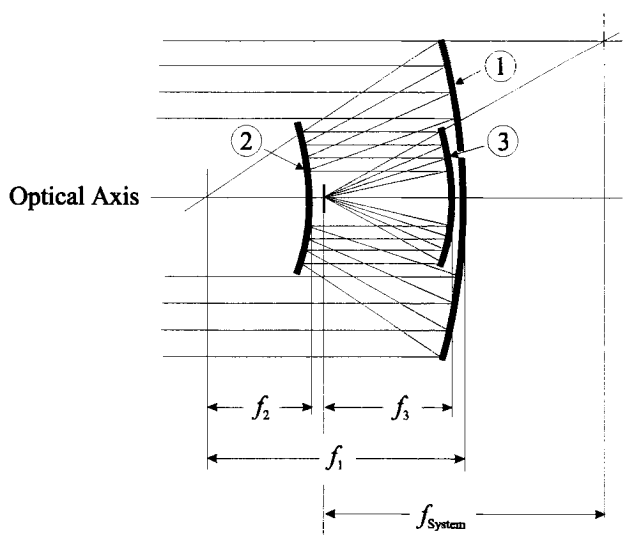


FIG. 5.28 Initial three-mirror system layout.

In a real system, this initial layout may be modified to use only off-axis portions of the individual elements as indicated in Fig. 5.29. This modification shows also how the primary and tertiary mirrors can be combined into a monolithic structure, suitable for single-point diamond turning. (The diamond turning process will be described in Chapter.10).

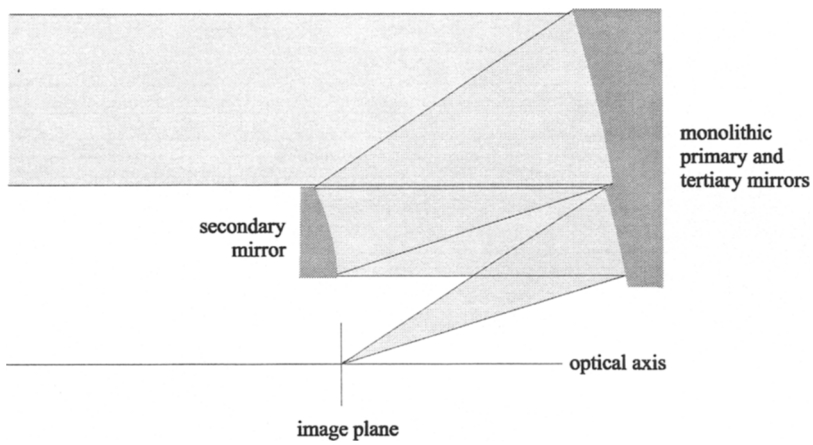


FIG. 5.29 Three-mirror system with monolithic primary and tertiary mirrors.

It is sometimes helpful to picture a mirror system as lenses. Figure 5.30 shows such a lens-equivalent layout of Fig. 5.29. It becomes clearer what the system's effective focal length is. Notice also the decentered aperture stop which indicates the partial use of the lens—or mirror elements, centered about the rotational axis.

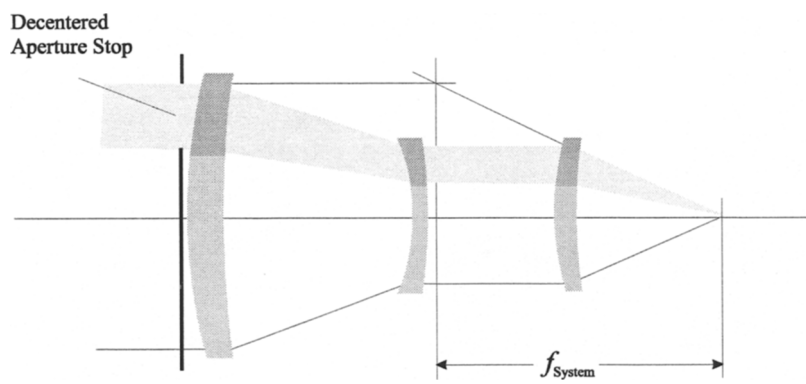


FIG. 5.30 Lens-equivalent configuration of three-mirror system show in Fig. 5.29.

Figures 5.31 and 5.32 show two additional three-mirror configurations, one as an objective, the other as an afocal system. Both present a clear demonstration of the beautiful simplicity of conic sections.

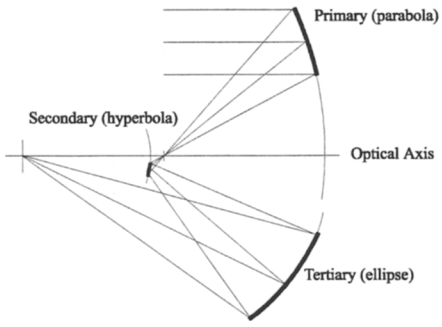


FIG. 5.31 Three-mirror objective system

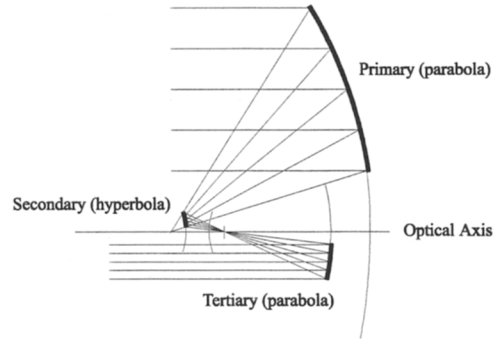


FIG. 5.32 Afocal three-mirror system

All this systems discussed so far have been only corrected for spherical aberration. To correct for other aberration requires the proper selection of separation, shape, and asphericity of the elements as well as the location of the aperture stop. One can see immediately that this is not a trivial task, and certainly beyond any fundamental introduction to lens design. The reader who is interested in studying the subject in depth is encouraged to consult references 5 and 6.

5.7 Diffractive (Binary) Optics

As a practical matter, one can look at diffraction simply as another method besides refraction and reflection to change the direction of light rays. Fig. 5.33 shows the three different ways of redirecting rays by an optical element.

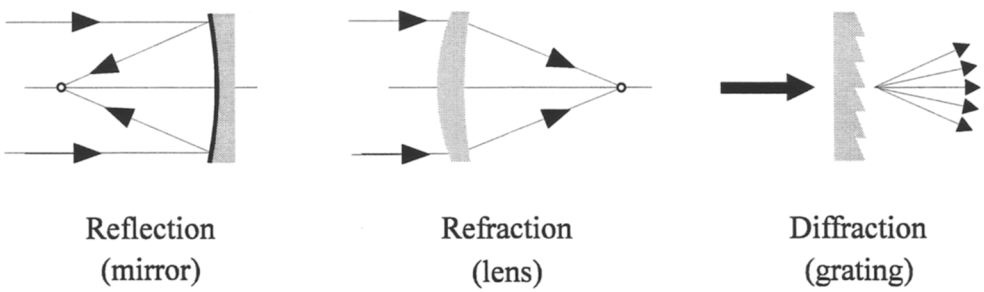


FIG. 5.33 Three ways to change direction of light rays.

Usually, one thinks first of a linear grating as a diffractive element, because it has been in use for a long time. By arranging a diffractive phase profile properly as concentric rings on a flat substrate, the grating becomes a diffractive lens. Such a lens is sometimes referred to as a single-wave Fresnel lens because of the one-wave optical path difference that exists at the zone transitions.

5.7.1 The simple diffractive singlet

A diffractive singlet as shown in Fig. 5.34 consists of a plane-parallel disk with a circular grating structure imposed on one surface. The performing principle of this

surface-relief structure lies in the periodic phase delay introduced by the continuous change of the physical thickness of the element according to the phase equation

$$\phi(r) = \frac{2\pi}{\lambda_0}\left(ar^2 + br^4 + cr^6 + \ldots\right). \tag{5.42}$$

In this equation, *a, b, c,* etc., are the phase coefficients, λ_0 is the wavelength and r is the radial coordinate. Diffractive zone boundaries occur at each 2π transition. The maximum zone depth at the transition is

$$d_{\max} = \frac{\lambda_0}{N_0 - 1}\ . \tag{5.43}$$

N_0 is the refractive index of the lens material at the design wavelength λ_0.

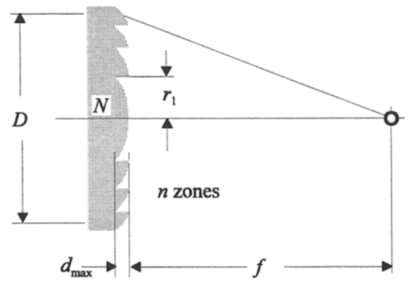


FIG. 5.34 Cross section of a simple diffractive lens (single-wave Fresnel lens).

The location of the zones in a diffractive lens can be determined from the equation for the classical Fresnel zone plate with

$$r_n = \sqrt{2n\lambda_0 f + n^2\lambda_0^2}\ . \tag{5.44}$$

If $f >> n\lambda_0$,

$$r_n \cong \sqrt{2n\lambda_0 f}\ , \tag{5.45}$$

where n is the zone number. Therefore, the first zone radius $r_1 \cong \sqrt{2\lambda_0 f}$ and $r_n \cong r_1\sqrt{n}$, from which we derive the total number of zones required for a lens with a diameter D.

$$n_{\text{total}} \cong \left(\frac{D}{2r_1}\right)^2 = \frac{f}{8\lambda_0(f/\#)^2}. \tag{5.46}$$

Equation (5.46) indicates that the required zone number for a diffractive lens is independent of the chosen substrate material.

Based on the quadratic term of the phase equation [Eq. (5.42)], $r_1 = \sqrt{\lambda_0/a}$, which leads to the relation $2\lambda_0 f = \lambda_0/a$, or

$$a = -\frac{1}{2f}. \tag{5.47}$$

The minus sign indicates the direction of the phase profile.

For a lens with a focal length of $f = 80$ mm and a diameter of $D = 16$ mm, used in the infrared spectrum at a wavelength of $\lambda_0 = 4$ μm, we find:

$a = -0.00625$ mm^{-1}, $r_1 = 0.8$ mm, and $n_{\text{total}} = 100$ zones.

$r_2 = 1.131$ mm
$r_3 = 1.386$ mm
...
$r_{99} = 7.960$ mm
$r_{100} = 8.00$ mm.

The zone spacing at the edge of the lens has diminished from the spacing of 0.8 mm for the first zone to 0.04 mm. This is an important indicator for the manufacturing process to be chosen.

As a quick reference approximation, the minimum zone spacing is stated by

$$(\Delta r)_{\text{min}} \cong 2\lambda_0 (f/\#). \tag{5.48}$$

If we select silicon, a suitable material for $\lambda_0 = 4$ μm, where its index $N_0 = 3.425$, the maximum zone depth $d_{\text{max}} = 1.65$ μm.

5.7.2 The hybrid achromat

For many applications, especially for broadband systems in the infrared region, the hybrid achromat presents a very interesting alternative to the conventional doublet.[7] By combining refractive and diffractive powers, one can effectively correct the chromatic aberration. The basic principle for this correction lies in the fact that a diffractive lens focuses the longer wavelengths closer to the lens than the shorter wavelengths. As indicated in Fig. 5.35, this is just the opposite of the behavior of a refractive lens. The total power of the refractive/diffractive combination is the sum of the powers of the elements:

$$\phi_{\text{hybrid}} = \phi = \phi_{\text{refractive}} + \phi_{\text{diffractive}}, \tag{5.49}$$

$$\text{or} \qquad f = \frac{f_r f_d}{f_r + f_d}, \qquad \text{with} \quad f = \frac{1}{\phi}. \tag{5.50}$$

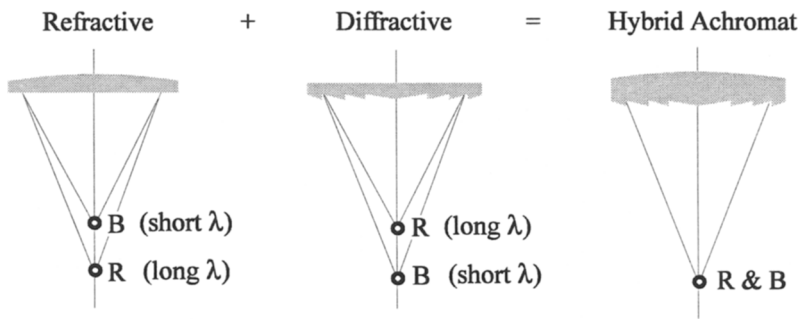


FIG. 5.35 Principle of combining refractive and diffractive powers for color correction.

By assigning the proper power to each element, their individual chromatic contributions cancel.

The following conditions must exist to achieve the correction:

$$f_r = \left(1 - \frac{\nu_d}{\nu_r}\right) f, \tag{5.51}$$

$$\text{and} \qquad f_d = \left(1 - \frac{\nu_r}{\nu_d}\right) f, \tag{5.52}$$

$$\text{with} \qquad \nu_r = \frac{N_0 - 1}{N_S - N_L} \qquad \text{and} \quad \nu_d = \frac{\lambda_0}{\lambda_S - \lambda_L},$$

where ν_r is the Abbe number for refraction, and ν_d is the Abbe number for diffraction. λ_0 is the center wavelength for which the index of refraction is N_0; and λ_S and λ_L are the short- and long-wavelength limits of the spectral band. Notice that ν_d is independent of any material used for the lens.

With these relationships, it becomes quite simple to estimate quickly, based on the first term (the quadratic term) of the phase equation [Eq. (5.42)], what the first zone radius will be and how many zones are needed for a particular lens.

$$r_1 = \sqrt{2\lambda_0 f_d}. \tag{5.53}$$

$$n_{\text{total}} = \left(\frac{D}{2r_1}\right)^2. \tag{5.54}$$

5.7.3 Numerical examples

To correct an $f/1$, 100-mm focal length germanium lens used in the 8–12-μm region, for which we select 10 μm as the center wavelength, we find that $v_r = 990$, $v_d = -2.5$, $f_r = 100.253$ mm, and $f_d = 39{,}700$ mm. Further, $r_1 = 28.178$ mm, $n_{total} = 4$, with $r_2 = 39.450$ mm, and $r_3 = 48.806$ mm. The fourth zone ends at the edge of the lens, where $r_{max} = 50$ mm. The maximum depth of the zones at their transitions is $d_{max} = 3.33$ μm.

From the equations we can see that large Abbe numbers and long wavelengths result in low zone numbers, as demonstrated with the example. This very favorable characteristic exists for most common materials used in the mid- and long-IR wavelength regions. A phase profile with only a few zones can be machined very economically by single-point diamond turning.[7] (See also Chapter 10)

For a quick reference, Figs. 5.36 and 5.37 show nomograms covering achromatic hybrids for the 3- to 5- and 8- to 12-μm windows. Only materials that can be diamond turned have been included in these nomograms.

In Fig. 5.36, a ZnSe lens with 50-mm focal length f is investigated. The first zone radius $r_1 = 6$ mm. For a lens diameter of $D = 50$ mm ($f/1$), the total zone number required to correct chromatic aberration is found to be $n_{total} \cong 17$. Table 5.1 lists all numerical values zone radii for the lens discussed.

Table 5.1 Zone radii for $f/1$ ZnSe example lens.

Zone n	Radius r	Zone n	Radius r
1	6.000 mm	10	18.974 mm
2	8.485	11	19.900
3	10.392	12	20.785
4	12.000	13	21.633
5	13.416	14	22.450
6	14.697	15	23.238
7	15.875	16	24.000
8	16.971	17	24.739
9	18.000	18	25.456 ($D/2 = 25$ mm)

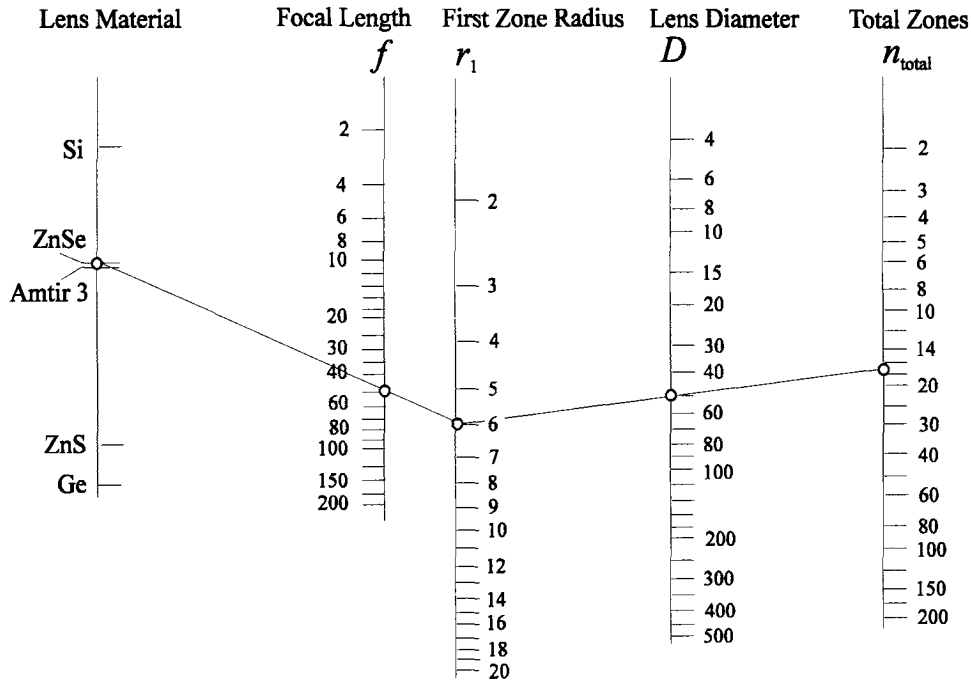


FIG. 5.36 Nomogram for hybrid achromat for the 3- to 5-μm region (object located at infinity).

This nomogram quickly assesses whether a proposed hybrid achromat requires a reasonable number of zones. For economic reasons, the tool radius is kept as large as possible in single-point diamond turning. However, the larger the tool radius, the more transmitted energy is being redirected at each zone transition step and turned into stray radiation. Fortunately, more than 20 zones are seldom needed for IR applications. The blockage or shadowing effects caused by the cutting tool can amount to several percent.[5] (See also Chapter 10).

For the 8- to 12-μm region, germanium is a superior material with its high Abbe number of almost 1000. It is in a class of its own compared to other diamond-turnable materials. This is indicated in Fig. 5.37 by the large separation from the other materials on the material scale. An $f/1$ germanium lens with 100-mm focal length requires only three steps (four zones) across its diameter. From $\beta_{chrom} \cong 1/[2v(f/\#)]$ one can see that for germanium, the angular chromatic blur amounts to only $0.0005/(f/\#)$ rad.

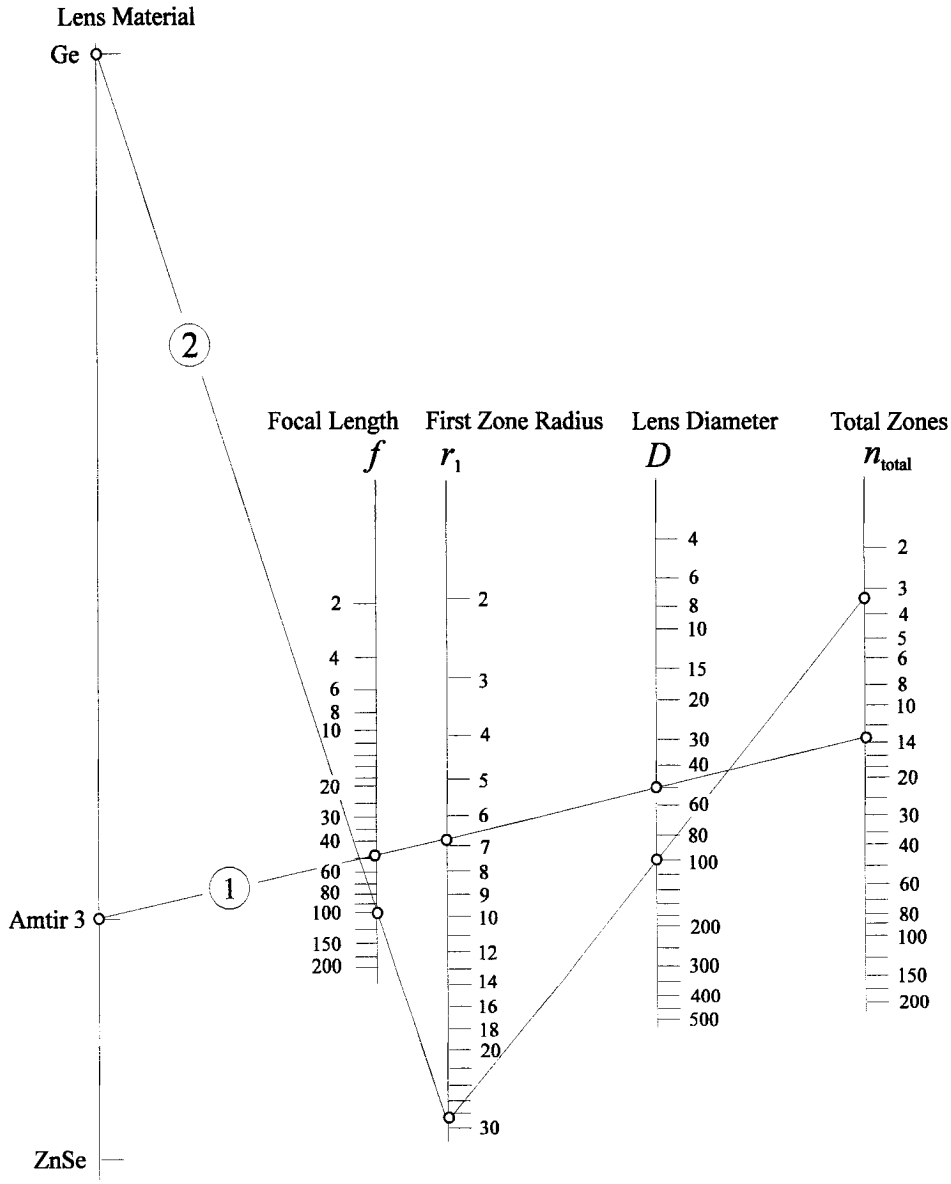


FIG. 5.37 Nomogram for hybrid achromat for the 8-to 12-μm region (object located at infinity).

In Fig. 5.37, example (1) shows an $f/1$ Amtir 3 lens with a 50-mm focal length. The first zone radius is found to be 6.7 mm, leading to a total of 14 zones across the lens diameter. The second example (2) is the $f/1$ germanium lens discussed above with a focal length of 100 mm. Here, the first radius is quite large, measuring 28.3 mm. Over the 100-mm diameter, only 4 zones are required to correct chromatic aberration.

5.7.4 Diffraction efficiency

As Eq. (5.53) indicates, the zone radii are calculated with reference to the center wavelength λ_0. At any other wavelength, detuning occurs. This means that the diffraction efficiency will be reduced.

Based on the scalar theory,[8] the diffraction efficiency for the first diffraction order is stated by

$$\varepsilon_1 = \left\{ \frac{\sin\left[\pi\left(\frac{\lambda_0}{\lambda} - 1\right)\right]}{\pi\left(\frac{\lambda_0}{\lambda} - 1\right)} \right\}^2 . \tag{5.55}$$

Fig. 5.38 shows the diffraction efficiency ε_1 as a function of the wavelengths ratio λ_0/λ. To balance the efficiency over a given spectral band, between λ_1 and λ_2 the center wavelength should be chosen as

$$\lambda_0 = \frac{2\lambda_1\lambda_2}{\lambda_1 + \lambda_2} . \tag{5.56}$$

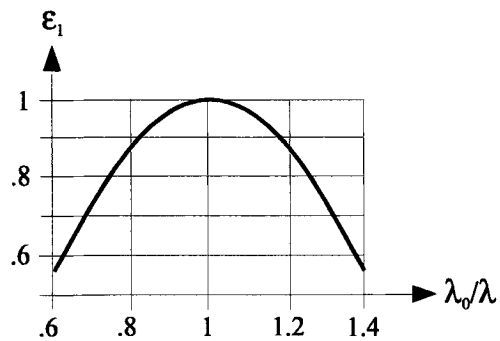


FIG. 5.38 Diffraction efficiency as a function of the ratio λ_0 and λ.

The approximate average efficiency can be stated by

$$\bar{\varepsilon}_1 \cong 1 - \frac{\pi^2}{36}\left[\frac{\Delta\lambda}{\lambda_0}\right]^2 , \tag{5.57}$$

where $\Delta\lambda = \lambda_2 - \lambda_1$.

As an example, for $\lambda_1 = 8$ and $\lambda_2 = 12$ μm, the center wavelength for equal efficiency reduction at both ends of the spectral window would be

$\lambda_0 = (2\times8\times12)/(8+12) = 9.6$ μm. With that, the average diffraction efficiency $\bar{\varepsilon}_1 \cong 95\%$. If we would have chosen 10 μm as the center wavelength instead of 9.6, the efficiency would have dropped to 81% at 8 μm and to 91% at 12 μm. The average efficiency would have stayed about the same. The shifting of the center or blaze wavelength can be effectively employed to favor one end of the spectral window relative to the other. This can be of interest for applications where signal strengths need to be balanced.

5.7.5 "Useful" spectral bandwidth

Rearranging Eq. (5.57) and solving for the bandwidth $\Delta\lambda$ yields

$$\Delta\lambda \cong \frac{6\lambda_0}{\pi}\sqrt{1-\bar{\varepsilon}_1}\,. \tag{5.58}$$

Assuming an average diffraction efficiency of 95%,

$$\Delta\lambda \cong 0.427\lambda_0\,. \tag{5.59}$$

For the LWIR region, with $\lambda_0 = 10$ μm, the "useful" bandwidth is therefore 4.27 μm. At $\lambda_0 = 4$ μm, the center of MWIR, $\Delta\lambda = 1.71$ μm, and for the visible spectrum with $\lambda_0 = 0.55$ μm, $\Delta\lambda = 0.23$ μm. These numbers reveal that over the common bandwidths, 5% or more of the radiation will be diffracted into other orders. In other words, there will be at least 5% stray radiation, which may cause some system problems. One can also look at this from a different perspective. The stray radiation will raise the background signal level and therefore lower the contrast. This affects the modulation transfer function (MTF).[9] MTF will be covered in Chapter 9.

The distribution of the energy into other than the first order will be discussed next, and an example for the MWIR region will be presented.

5.7.6 Diffraction efficiency for a particular order

With some assumptions and simplifications, the diffraction efficiency ε_m for a particular order m can be stated as

$$\varepsilon_m = \frac{\sin^2\left[\pi\left(\frac{\lambda_0}{\lambda} - m\right)\right]}{\left[\pi\left(\frac{\lambda_0}{\lambda} - m\right)\right]^2}\,. \tag{5.60}$$

Figure 5.39 shows the diffraction efficiency of several orders for a diffractive element designed for the 3- to 5-μm region. Notice that the drop at the limiting wavelengths is equal for the first order, which results from the chosen design wavelength $\lambda_0 = 3.75$ μm. [See the remarks on page 113 and Eq. (5.56)].

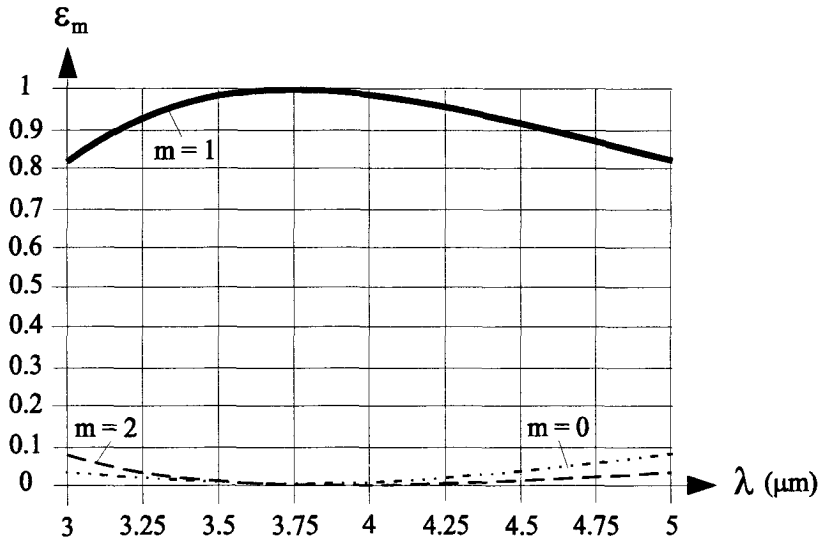


FIG. 5.39 Diffraction efficiency for the zero, first, and second orders in the MWIR region.

5.7.7 The hybrid achromat, corrected for chromatic and spherical aberrations

Changing one of the two surfaces of a singlet to an asphere will correct the lens for spherical aberration. To eliminate chromatic aberration we add the diffractive phase profile as described above. A good choice is to combine the two approaches by superimposing the phase profile onto the aspheric surface. For better protection against the environment, the second lens surface is preferred.

We remember from Eqs. (3.5) and (3.6) that a thin lens shaped for minimum spherical aberration has the following radii:

$$R_1 = \frac{2(N+2)(N-1)}{N(2N+1)} f \quad \text{and} \quad R_2 = \frac{2(N+2)(N-1)}{N(2N-1)-4} f .$$

The aspheric shape (conic section for third-order correction) of the second surface for such a basic lens is expressed by the conic constant

$$p_2 = 1 + \frac{2N(4N-1)(N+2)^2}{[N(2N-1)-4]^3}. \tag{5.61}$$

For comparison, the conic constants for the four different basic materials mentioned frequently throughout this tutorial text are

Material	N	p_2	**Shape**
Glass	1.5	−182.75	Hyperbola
Zinc Selenide	2.4	6.9544	Ellipse
Silicon	3.4	1.6431	Ellipse
Germanium	4.0	1.3125	Ellipse

Applying these equations to the *f*/1 germanium singlet with a focal length of 100 mm, used for the 8- to 12-μm window, we correct for spherical and chromatic aberrations and find:

First radius	$R_1 = 100$ mm	[Eq. (3.5)]
Second radius	$R_2 = 150$ mm	[Eq. (3.6)]
Conic constant	$p_2 = 1.3125$	[Eq. (5.61)]
First zone radius	$r_1 = 28.2$ mm	[Fig. 5.36 or Eq. (5.45)]
Total zones required	$n_{\text{total}} = 3.14 = 4$	[Fig. 5.36 or Eq. (5.46)].

Using these numbers as a starting point, adding a suitable thickness and the refined, real index of refraction, and optimizing the parameters with a computer yields in a matter of seconds an achromatic hybrid that does not vary too much from this thin-lens third-order-aberration-corrected singlet, even for the very low relative aperture of *f*/1. This is another verification of the thin-lens approach for the IR.

5.7.8 Binary optics

The term *binary optics* relates to the manufacturing process of a diffractive surface. In the photolithographic process the phase profile is not etched into the substrate material as a continuous surface within a given zone, but approximated with a number of steps. This is illustrated in Fig. 5.40. The relationship between the number M of photolitographic masks needed to produce K steps or levels is

$$K = 2^M \quad . \tag{5.62}$$

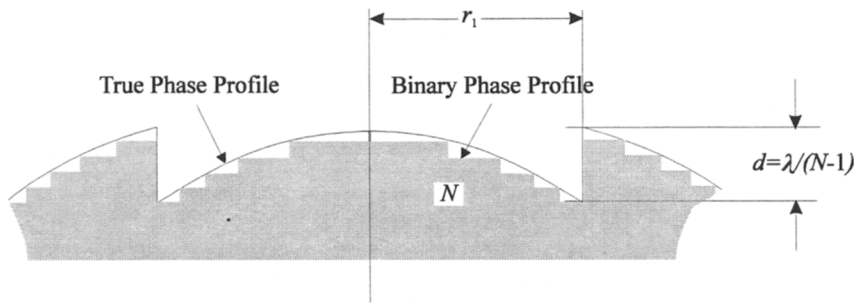


FIG. 5.40 Binary optics approximates the true phase profile.

The diffraction efficiency of such a binary structure is determined by

$$\varepsilon_1 = \left[\frac{\sin(\pi/K)}{(\pi/K)} \right]^2 \quad . \tag{5.63}$$

From this equation it can be seen that it takes 16 steps or levels to achieve a 99% diffraction efficiency. The number of masks required to produce 16 steps is 4 ($2^4 = 16$). Considering that the total zone depth is only 3.33 μm for a germanium lens applied at a wavelength of 10 μm, the individual step for the binary lens profile is $3.33/16 \cong 0.208$ μm. This requires precise control in the etching process. Moreover, the lateral alignment of the masks has to be accurate as well to achieve concentricity of the zones relative to the center of the lens. It becomes a very demanding manufacturing challenge to add the binary phase profile to a curved surface. In contrast, it is convenient to just add the diffractive phase function to the function of the base surface when the lens material can be machined with a single-point diamond tool.

An interesting and powerful application of binary optical elements is their use as field lenslets or even immersion lenslets to achieve an optical gain with focal plane arrays. Figure 5.41 indicates such an application. The microlenslet images the aperture stop onto the detector element. This has an integrating effect over the detector element. More importantly, by reducing the detector element size, space is gained around the elements without reducing the effective fill factor of the array. This space can be used advantageously for electrical connections of the elements. The arrangement is shown as an assembly of a micro-optics array and a detector array. In principle, the two can be combined into one component by placing the binary lenslets on the front side of the substrate and the detector array on the back side. The success of such an arrangement is strictly a matter of manufacturing yields. To make the challenge even greater, one can consider depositing spectral filters directly onto the optical side of the array.

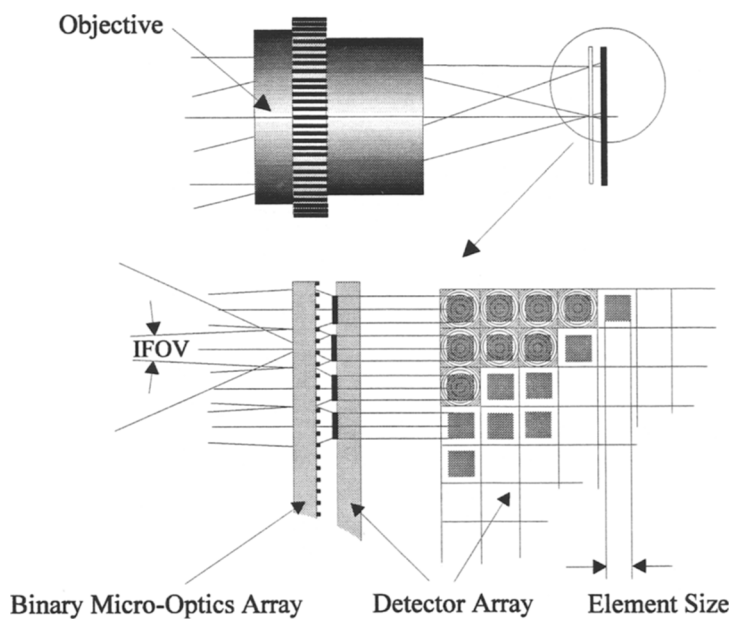


FIG 5.41 Binary micro-optics applied to a focal plane array.

References

1. H. Dubner, "Optical Design for Infrared Missile-Seekers," Proceedings of the IRE (September 1959), pages 1537-1539.
2. W. J. Smith, *Modern Optical Engineering*, McGraw-Hill, Inc. (2000), page 286.
3. R. Kingslake, *Lens Design Fundamentals*, Academic Press (1978), pages 112 and 113.
4. W. J. Smith, *Modern Optical Engineering*, McGraw-Hill, Inc. (2000), pages 479-481.
5. D. Korsch, *Reflective Optics*, Academic Press, Inc. (1991).
6. L.CG. Cook, "The last three-mirror anastigmat (TMA)?," SPIE Critical Review, Vol. CR 41 (1992), pages 310–324.
7. M. J. Riedl, "Predesign of diamond turned defractive/diffractive elements for IR objectives," SPIE Critical Review, Vol. CR 41 (1992), pages 140–156.
8. G. J. Swanson, *Binary Optics Technology: The Theory and Design of Multilevel Diffractive Elements*, Technical Report 854, Lincoln Laboratory, (1989).
9. D. A. Buralli and G. M. Morris, "Effects of diffraction efficiency on the modulation transfer function of diffractive lenses," Applied Optics, Vol. 31, No. 22 (Aug. 1992).

CHAPTER 6

Design Examples

6.1 Introduction

After having covered the aberration effects, we can go back to the S/N calculations we performed in Section 2.8.1 for high- and low-temperature applications. We shall analyze several optical configurations and determine the impact of the aberrations for those chosen arrangements. As an additional design example, the Petzval objective from Chapter 3 will be modified to improve its performance by using an aspheric diffractive element. At the end of the chapter, some remarks about instantaneous field of view (IFOV) are presented.

6.2 Basic Assumptions for the High- and Low-Temperature Applications

We assume that the field is scanned, using a 45-element linear detector array. Each detector element measures 0.5 mm square. With a spacing of 0.1 mm between each of the elements, the pitch of the array is 0.6 mm. The basic optical system layout with a single lens is shown in Fig. 6.1.

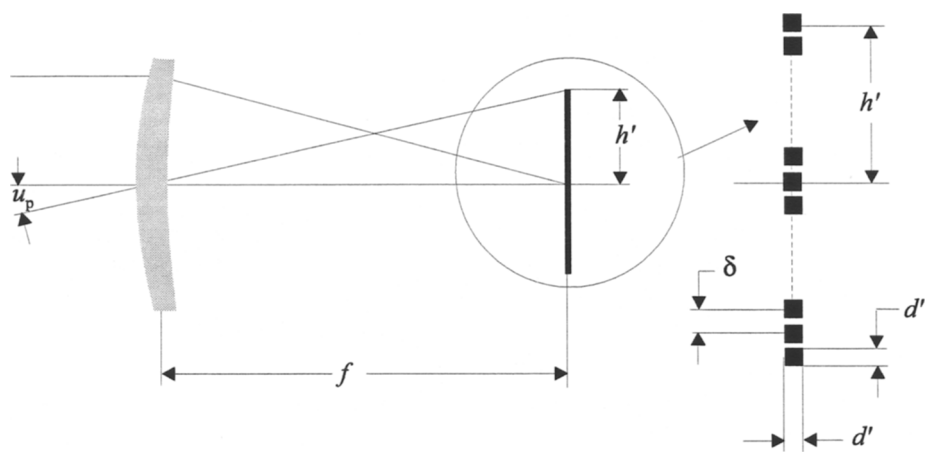


FIG. 6.1 Basic optical system parameters: $f = 200$ mm, $d' = 0.5$ mm, $\delta = 0.6$ mm, and $n = 45$ elements.

With these guideline, we can calculate

$$h' = 0.5\,(n-1)\,\delta = 13.2 \text{ mm, and } u_p = h'/f \approx 0.07 \text{ rad}.$$

6.2.1 Optics for high-temperature systems (3–5 μm)

The 200-mm-focal-length singlet is made from silicon with $N = 3.4$ and $v = 235$. Its relative aperture is $f/4$.

Applying the thin-lens third-order expression from Chapter 3, we find that for a lens shaped for minimum spherical aberration $R_1 = 195.48$ mm and $R_2 = 329.77$ mm. The aberrations add up to a total blur spot size of $B \cong 0.273$ mm.

To demonstrate the validity of the thin-lens third-order aberration concept, a thickness of 5 mm was added and the lens was optimized with a computer program. The results are shown in Fig. 6.2.

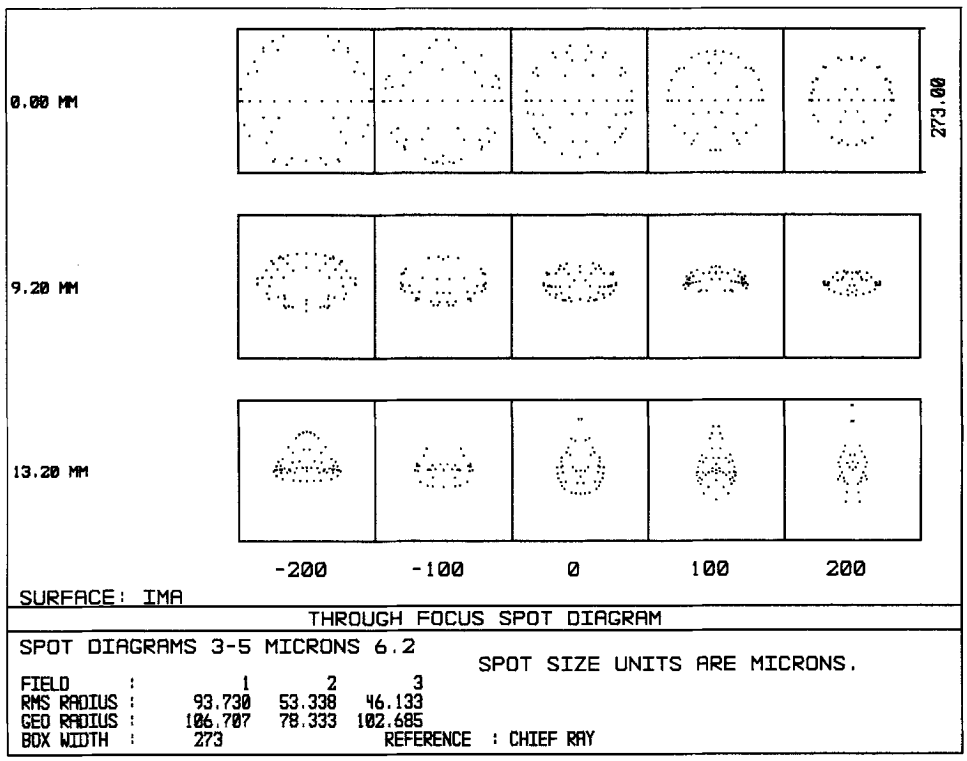


FIG. 6.2 Spot diagrams for optimized silicon singlet.

The scales in the spot diagrams have been matched with the predicted thin-lens third-order results for direct comparison (The box width is equal to blur spot dimension 0.273 mm.) To balance the energy distribution over the full field, the concept of field curvature compensation as shown in Fig. 3.19 was applied.

6.2.2 Optics for low-temperature systems, design #1 (8–12 μm)

The lens material for the singlet used for this application is germanium with an index of $N = 4$ and an inverse relative dispersion of $v = 990$. The relative aperture is $f/2$.

The thin-lens third-order results for this case are R_1 = 200 mm, R_2 = 300 mm, and the blur spot size $B \cong 0.55$ mm. As expected, the computer-optimized results also agree here very well. We notice that the blur spot is larger than the 0.5-mm detector element size. Therefore, we look at another configuration, the Petzval objective.

6.2.3 Optics for low-temperature systems, design #2 (8–12 μm)

Applying Eqs. (3.34) and (3.35) for an $f/2$ Petzval objective yields an on-axis blur due to spherical and chromatic aberrations of $B \cong 0.062$ mm. Adding coma, astigmatism, and lateral color for the 4° field angle brings the off-axis blur to approximately 0.37 mm. The major contributor is astigmatism, with 0.2 mm. Optimizing the radii and the image plane position results in additional improvements. This is indicated in Fig. 6.3.

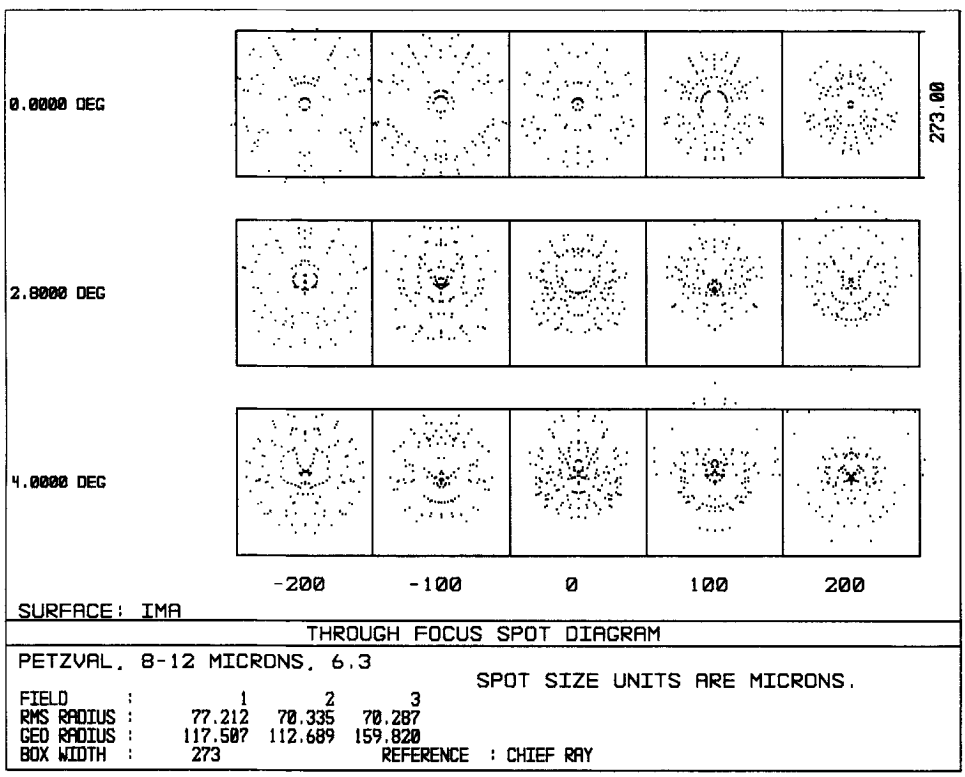


FIG. 6.3 Spot diagrams for the germanium Petzval objective.

6.2.4 Optics for low-temperature systems, design #3 (8–12μm)

As an interesting exercise, we shall analyze an optical system that consists of two aspheric mirrors and an aspheric corrector. This configuration is pictured in Fig. 5.27. First, we make some reasonable assumptions, which provide the basic dimensions *e, d,* and *e.* With a 25% central area obstruction, *b* becomes $0.5f$ or 100 mm (we remember that f = 200 mm and $f/\# = 2$). Not exceeding $f/1$ as relative aperture for the primary mirror (for manufacturing reasons) brings a spacing *d* of 50 mm. Deciding that the corrector will serve as the Dewar window for the detector package, we place it 25 mm in front of the

focal plane, i.e., $e = 25$ mm. Finally, as material for the corrector, we use germanium with index $N = 4$. From the basic Cassegrain Eqs. (3.57) and (3.58), we find the radii for the mirrors. They are identical: $R_1 = R_2 = -200$ mm. Equations (5.39) through (5.41) are used to determine the conic constants for the mirrors and the deformation coefficient for the corrector plate:

$$p_1 = -0.769231, \quad p_2 = -20.66667, \quad A_{a_3} = -9.57265 \times 10^{-7}.$$

Both mirrors are hyperbolas. It turns out that the Petzval contribution, which determines the field curvature, is zero for the selected dimensions. That means the image surface is flat.

To correct for higher-order aberrations, two deformation coefficients were added to the basic conic sections. The final result is shown in Fig. 6.4.

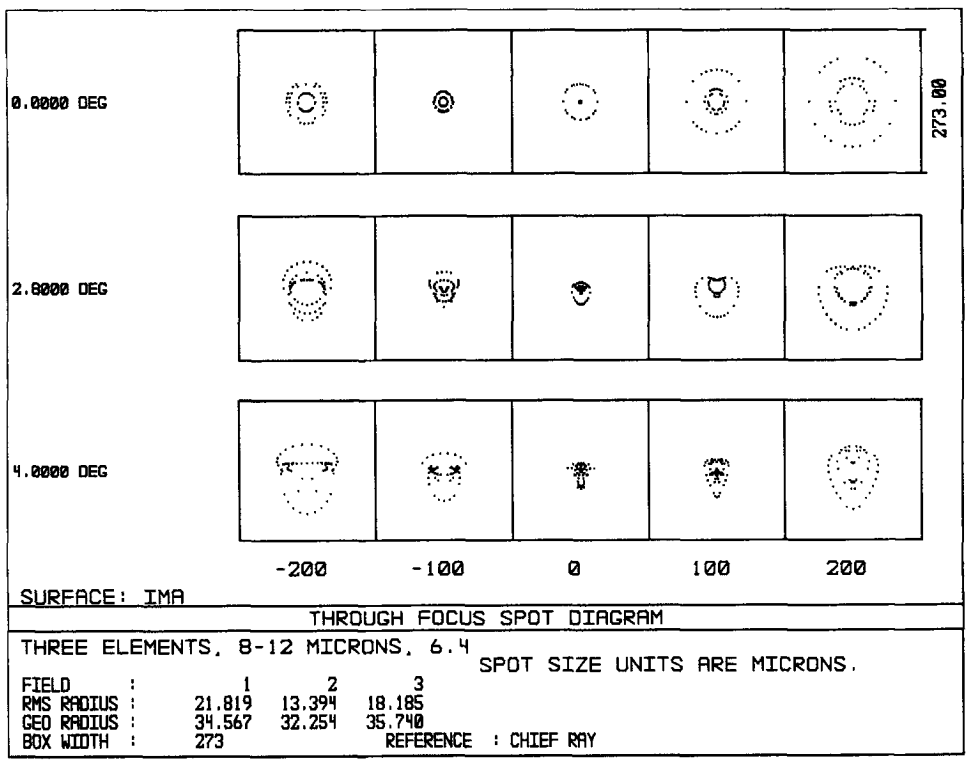


FIG. 6.4 Spot diagram for the three-element system.

As mentioned before, for better comparison, the box size surrounding the blur spots has been kept the same for all cases shown, namely 0.273 mm, the size of the blur spot for the $f/4$ singlet. It should be noted that the central obstruction reduces the amount of the transmitted radiation by the ratio of the area of the secondary mirror to the area of the primary mirror. One can express that as an effective ($f/\#$) by

$$(f/\#)_{\text{effective}} = \frac{1}{\sqrt{1-\varepsilon^2}}(f/\#), \tag{6.1}$$

where ε is the ratio of the secondary and primary mirror diameters. In the case discussed $\varepsilon = 0.5$, which yields an effective $(f/\#)$ of 2.3. It is easy to see that the S/N will be reduced by 25%. If this is not acceptable, the initial $(f/\#)$ has to be lowered accordingly and the impact on the aberrations has to be reevaluated.

6.3 The Improved Petzval Objective

This example will serve several purposes. First, it will teach how the material presented so far can be effectively used in the predesign stages; and second, it will demonstrate how such a fundamental approach can provide a better understanding of the interaction among the different parameters. Finally, it will also indicate that arriving at a suitable optical configuration to meet the existing requirements is not a process often called "automatic lens design."

We begin with the basic configuration as described in Chapter 3, especially as shown in Fig. 3.23.

To present the entire design process, we repeat the equations for the radii:

$$R_1 = \frac{4(N+2)(N-1)}{(2N+1)N} f, \tag{6.2}$$

$$R_2 = \frac{4(N+2)(N-1)}{N(2N-1)-4} f, \tag{6.3}$$

$$R_3 = \frac{2(N+2)(N-1)}{N(6N+1)-4} f, \tag{6.4}$$

and

$$R_4 = \frac{2(N+2)(N-1)}{N(6N-1)-8} f. \tag{6.5}$$

The angular blur spot, due to spherical aberration is, as has been already stated,

$$\beta_{\text{spher}} = \frac{N\left[4N(2N-7)+11\right]}{2{,}048(N-1)^2(N+2)(f/\#)^3}. \tag{6.6}$$

To correct this residual spherical aberration, we apply the procedure described in Ref. 1 and find that if the second surface of the front element is changed to a conic section, its conic constant should be[2]

$$p = 1 - \frac{1}{96}\left[\frac{4(N+2)(N-1)}{N(2N-1)-4}\right]^3. \tag{6.7}$$

The axial chromatic aberration of the arrangement under discussion is

$$\beta_{chrom} = \frac{3}{8\nu_r (f/\#)}, \tag{6.8}$$

where ν_r is the refractive Abbe number.

The equation for the diffractive phase profile used to reduce the axial chromatic aberration can be derived from the general phase equation [Eq. (5.42)]. Limiting the expression to the quadratic term yields for the chosen Petzval configuration

$$\phi(r) = \frac{2\pi}{\lambda_0} a r^2, \tag{6.9}$$

with the quadratic phase constant $$a = -\frac{3\nu_d}{8(\nu_d - \nu_r) f}. \tag{6.10}$$

ν_d is the diffractive Abbe number that is expressed by the ratio of the design wavelength and the difference between the short and long wavelengths.

6.3.1. Numerical example for a LWIR application

To demonstrate the usefulness of all these equations, we shall design a Petzval objective for the 8–12 μm region. By assuming an index of refraction of 4 and an Abbe number of 1000, the calculations become very simple. For an objective with a focal length of 100 mm and a relative aperture ($f/\#$) of 1, we find $R_1 = 2f = 200$ mm, $R_2 = 3f = 300$ mm, $R_3 = (3/8)f = 37.5$ mm, and $R_4 = (3/7)f = 42.857$ mm.

The conic constant p for the aspheric second surface of the front element becomes [with Eq. (6.7)]

$$p = 1 - \frac{1}{96}\left[\frac{4(4+2)(4-1)}{4(2\times 4-1)-4}\right]^3 = 1 - \frac{1}{96}\left[\frac{72}{24}\right]^3 = 1 - \frac{27}{96} = \frac{23}{32} = 0.71875,$$

which represents a prolate ellipsoid. The phase coefficient a is obtained with Eq. (6.10). It is

$$a = -\frac{3\times(-2.5)}{8(-2.5-1000)\times 100} = -9.35\times 10^{-6}.$$

The first radius of the phase profile is

$$r_1 = \sqrt{\frac{\lambda_0}{a}} = \sqrt{\frac{0.01}{9.35\times10^{-6}}} = 32.70 \text{ mm},$$

and the total number of zones [Eq. (5.54)] is

$$n_{z\,total} = \left(\frac{100}{2\times32.70}\right)^2 = 2.33,$$

which means the profile has 2 full zones. The third one does not reach its full step size d, because its radius is larger than the half-diameter of the lens. The second and third zone radii are

$$r_2 = r_1\sqrt{2} = 46.25\,\text{mm}, \quad \text{and} \quad r_3 = r_1\sqrt{3} = 56.64\,\text{mm}.$$

This step size is

$$d = \frac{0.01}{(4-1)} = 0.00333\,\text{mm} = 3.33\,\mu\text{m}.$$

The average diffraction efficiency amounts to $\bar{\varepsilon} \cong 1 - \frac{\pi^2}{36}\left(\frac{4}{10}\right)^2 \cong 0.956.$

Fig. 6.5 shows the spot diagram of this improved design. While the on-axis blur spot for the all-sphere arrangement in Section 3.6.1 was 135 μm in diameter, adding a diffractive asphere yields a diffraction limited system, even for the off-axis positions. Notice that the circle surrounding the spot diagrams is the Airy disk, which is 24.4 μm.

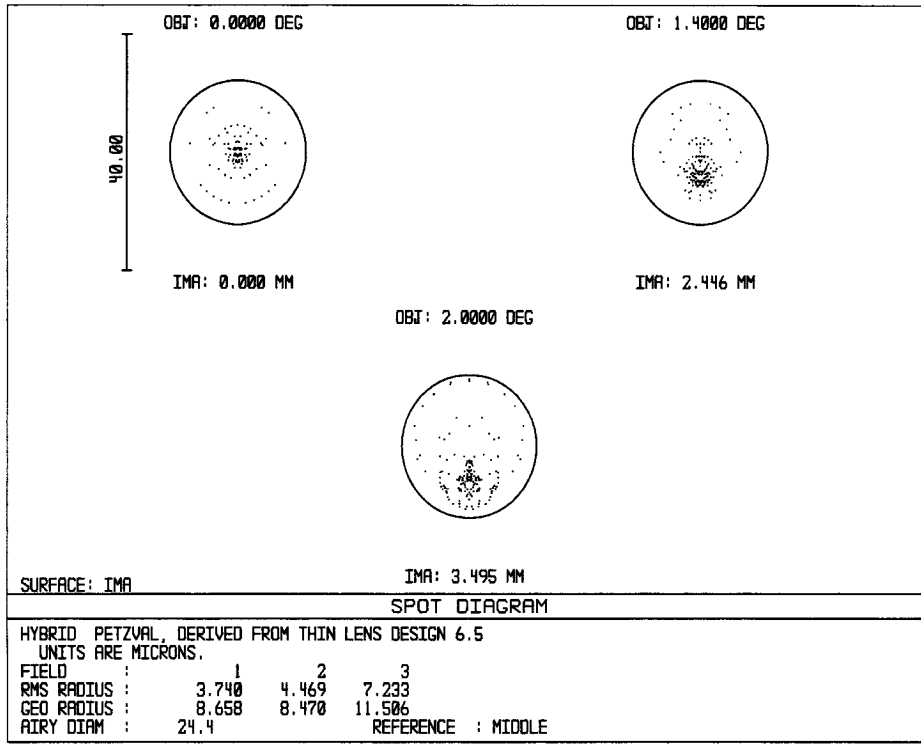


FIG. 6.5 Spot diagrams for the improved Petzval objective (hybrid), with a conic section surface and a superimposed diffractive phase profile on the rear surface of the front element.

6.3.2 Manufacturing remarks

Diamond turning is the preferred machining process to generate the aspheric surface upon which the diffractive phase profile is superimposed. By properly programming the motion of the single point diamond tool, both features can be produced simultaneously. However, there is an aspect to this process which can be limiting to its usefulness, especially if the spacing between the zones of the phase profile is narrow. This will be discussed in detail in Chapter 10.

6.4 Instantaneous Field of View

For a perfect optical system, the instantaneous field of view (IFOV) is strictly determined by the size of the detector element d'. Its angular extent is expressed as IFOV = $2u_p = d'/f$. The blur spot is at best the diffraction size of $B = 2.44\ \lambda f/\#$), but frequently larger, depending on the residual aberrations. In either case, one has to agree with the user on the energy percentage of the blur spot size that is acceptable for the application as a field cutoff.

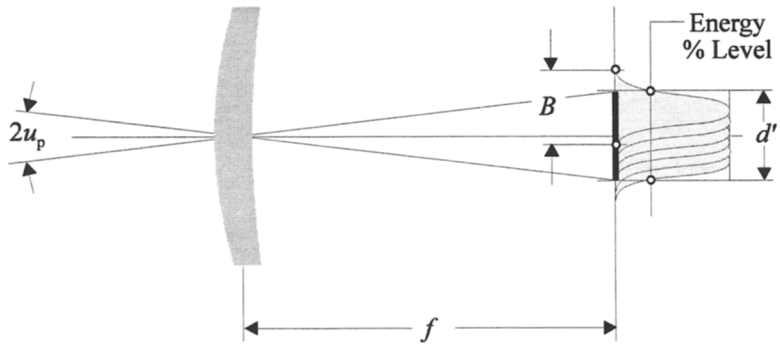


FIG. 6. 6 Blur spot and instantaneous field of view.

References

1. M. J. Riedl, "Predesign of diamond turned refractive/diffractive elements for IR objectives," SPIE Critical Review, Vol. CR 41 (1992), pp. 140–156.
2. W. J. Smith, *Modern Optical Engineering*, McGraw-Hill, Inc. (2000), pp. 332–335.

CHAPTER 7

Thermal Effects

7.1 Introduction

Frequently, optical systems must perform over a wide temperature range. Due to the thermal expansion and the change of the index of refraction of the lens material with temperature, the performance of the system is affected. This is especially pronounced in the infrared region, where most materials suffer from a high *dN/dT*, the change of index with temperature. The values for the index of refraction and especially the thermal coefficients vary—in some cases widely—in the literature and also in the data sheets from suppliers. It is therefore necessary to confirm or recheck these data.

To maintain an acceptable image quality, in many cases refocusing will be necessary. This can be accomplished mechanically or optically. Specifically, the mechanical adjustment can be done manually or by other means, such as feedback servo systems and others.[1] Optically, the compensation can be achieved by selecting suitable optical materials and element powers. This latter method will be discussed with the assumption that the lens system is stabilized at a given temperature; in other words, there are no temperature gradients present.

To close out the chapter, some remarks are made relating to the cold stop and cold shield.

7.2 Changing Parameters

The changing parameters of a lens in a housing are the following:

	at temperature t	at temperature $t+\Delta t$
1. Lens		
radius	r	$r(1+\alpha_L \Delta t)$
thickness	d	$d(1+\alpha_L \Delta t)$
index of refraction	N	$N+(dN/dt)\Delta t$
2. Housing (mount)		
spacing between lens and detector	s	$s(1+\alpha_H \Delta t)$

α_H = coefficient of linear expansion for housing

α_L = coefficient of linear expansion for lens

dN/dt = change of refractive index with temperature

The symbols are identified in Fig. 7.1:

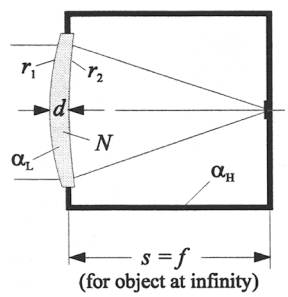


FIG. 7.1 Identification of changing parameters for a lens mounted in a housing.

7.3 Defocus with Change of Temperature

To simplify the procedure, thin elements are assumed. For an objective consisting of j thin elements mounted in a housing, the focal shift is

$$\Delta f = -\left[f^2 \sum_{i=1}^{j} \left(\frac{T_i}{f_i} \right) + \alpha_H f \right] \Delta t , \tag{7.1}$$

where

$$T_i = \frac{dN_i / dt}{(N_i - 1)} - \alpha_{L_i} = \text{thermal glass constant},$$

f = focal length of system (assumed to be the length of the housing),
f_i = focal length of element i,
N_i = refractive index of element i,
dN_i/dt = index change of element i,
a_{L_i} = thermal expansion coefficient of element i,
a_H = Thermal expansion coefficient of housing.

7.4 Defocus of Singlet

The defocusing distance Δf is the separation of the lens focus from the detector. This is depicted in Fig. 7.2, where a temperature increase is assumed. It is indicated that the focal length actually decreases with an increase in temperature, while the housing expands so that the total defocusing is the sum of both effects. This total separation of the lens focus from the detector position is

$$\Delta f = -f \Delta t \left(T + \alpha_H \right). \tag{7.2}$$

Some of the most frequently used IR lens materials with their respective glass constants are called out in Table 7.1. BK7, the very common lens material for the visible spectrum, is included for reference. Notice that its glass constant is negative, which means that the focal length for a lens made from BK7 increases with temperature. This is the

opposite from IR lenses made from suitable materials for that region, such as germanium, silicon, zinc sulfide, and others. Typical housing materials such as aluminum and steel with their thermal coefficients of expansion are also included in Table 7.1.

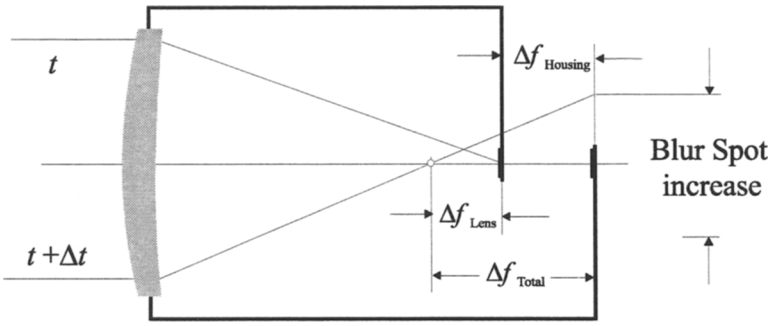


FIG. 7.2 Components of focus shift.

Table 7.1 Glass constants and coefficients of expansion for some frequently used materials.

Lens Material	Glass Constant $T \times 10^{-6}$	Housing Material	Coeff. of Expansion $\alpha_H \times 10^{-6}$ mm/mm/°C
BK7	-1 VIS	Aluminum	24
ZnS	25 IR	Steel 1015	12
ZnSe	36 IR	Invar 36	1.3
Si	60 IR		
Ge	126 IR		

We shall analyze with Eq. (7.2) the thermal effects for a germanium singlet with a focal length of 100 mm mounted in an aluminum housing. The temperature excursion is assumed to be +20°C.

$$\Delta f = -100\times20\times(126+24)\times10^{-6} = \underset{\Delta f_{\text{Lens}}}{-0.252}\underset{\Delta f_{\text{Housing}}}{-0.048} = \underset{\Delta f_{\text{Total}}}{-0.30}\ \text{mm}.$$

This little exercise shows how devastating a change in temperature can be without any correction for the image position, if one considers that the image blur spot increases according to $\Delta B = \Delta f_{\text{Total}}/(f/\#)$.

7.5 Athermalization with a Doublet

For two elements, Eq. (7.1) simplifies to

$$\Delta f = -f\left[f\left(\frac{T_A}{f_A}+\frac{T_B}{f_B}\right)+\alpha_H\right]\Delta t\,. \tag{7.3}$$

Element A is the front element and B is the rear element. To eliminate any focal shift, we set $\Delta f = 0$. For that case, the focal length of the front element is

$$f_A = \frac{\left(T_B - T_A\right)}{\left(T_B + \alpha_H\right)} f \,, \tag{7.4}$$

and the focal length of the rear element is

$$f_B = \frac{\left(T_A - T_B\right)}{\left(T_A + \alpha_H\right)} f \quad \text{or} \quad \frac{f_A f}{\left(f_A - f\right)}. \tag{7.5}$$

Applying these expressions for a doublet mounted in an aluminum housing with a silicon front element and a germanium rear element leads to a focal length of 44 mm for the silicon lens, and −78.6 mm for the germanium lens. The design wavelength has been chosen to be 4 μm. If we select a convex-plano front element and a plano-concave rear element, the respective radii are $r_1 = 105.6$ mm, $r_2 = \infty$, $r_3 = \infty$, and $r_4 = -235.8$ mm. This is a good starting point for a computer optimization after adding 3 and 2.5 mm for lens thicknesses, and a spacing of 1 mm between the elements. It is important to remember that while optimizing the radii to reduce spherical aberration, the focal lengths of the elements must be maintained, otherwise the athermalization effect will be lost. Figure 7.3 shows the expected blur spot sizes at −20°C and at +60°C for an f/4 doublet, designed and optimized for 20°C. The diffraction blur spot size is shown for reference.

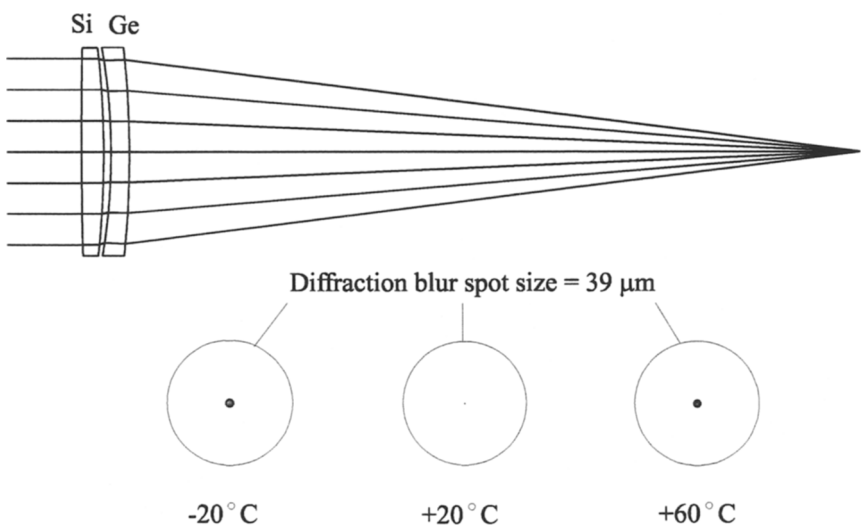


FIG. 7.3 Athermalized doublet (Athermat). The blur spots indicate that the system remains diffraction limited over a temperature excursion of ± 40°C.

7.6 The Athermalized Achromat

The equations for the doublet analyzed in the previous section are valid for one wavelength only. Such a doublet would therefore be suitable for laser applications or other monochromatic functions. To achromatize and athermalize an objective, three conditions must be satisfied. They are

$$\phi = \sum_{i=1}^{j} \phi_i , \tag{7.6}$$

$$\sum_{i=1}^{j} \left(T_i \phi_i \right) = -\alpha_H \phi , \tag{7.7}$$

and

$$\sum_{i=1}^{j} \left(\frac{\phi_i}{\nu_i} \right) = 0 . \tag{7.8}$$

As previously stated, ν is the Abbe number. For refractive elements, $\nu_r = \dfrac{N_M - 1}{N_S - N_L}$, and for diffractive elements, $\nu_d = \dfrac{\lambda_M}{\lambda_S - \lambda_L}$. The roots for Eqs. (7.6)– (7.8) are

$$1/f_2 = \phi_2 = \frac{\dfrac{1/\nu_1}{1/\nu_3 - 1/\nu_1} - \dfrac{T_1 + \alpha_H}{T_3 - T_1}}{\dfrac{1/\nu_1 - 1/\nu_2}{1/\nu_3 - 1/\nu_1} - \dfrac{T_1 - T_2}{T_3 - T_1}} \phi , \tag{7.9}$$

$$1/f_3 = \phi_3 = \frac{1/\nu_1 - 1/\nu_2}{1/\nu_3 - 1/\nu_1} \phi_2 - \frac{1/\nu_1}{1/\nu_3 - 1/\nu_1} \phi , \tag{7.10}$$

$$1/f_1 = \phi_1 = \phi - \phi_2 - \phi_3 . \tag{7.11}$$

For refractive elements, $T_r = \dfrac{dN/dt}{(N-1)} - \alpha_L$, and for diffractive elements, $T_d = 2\alpha_L$.

To apply these seemingly cumbersome equations to a practical example, we extend the design of the doublet from the previous section to cover the MWIR region from 3 to 5 μm. For the first design, three lens elements will be used. In the second design, the use of a diffractive element for an athermat will be demonstrated. As before, the effective focal length is 100 mm and the relative aperture is 4.

7.6.1 The all-refractive athermal achromat

As lens material, we elect for

Element 1: silicon with $\nu_1 = 236$, and $T_1 = 60\times10^{-6}$ (see Table 7.1)
Element 2: germanium with $\nu_2 = 103$, and $T_2 = 126\times10^{-6}$
Element 3: zinc sulfide with $\nu_3 = 110$, and $T_3 = 25\times10^{-6}$.

The housing is made from aluminum with $\alpha_H = 24\times10^{-6}$. Eqs (7.1), (7.2), and (7.3) result in

$$\phi_2 = \frac{\dfrac{1/236}{1/110-1/236} - \dfrac{60+24}{25-60}}{\dfrac{1/236-1/103}{1/110-1/236} - \dfrac{60-126}{25-60}} \times 0.01 = \frac{3.2730159}{-3.0130066} \times 0.01 = -0.0108630\ ,$$

$$\phi_3 = \frac{1/236-1/103}{1/110-1/236} \times (-0.0108630) - \frac{1/236}{1/110-1/236} \times 0.01 = 0.0035156\ ,$$

$$\phi_1 = 0.01-(-0.010863)-0.0035156 = 0.0173474\,.$$

The respective focal lengths are $f_1 = 57.646$ mm, $f_2 = -92.056$ mm, and $f_3 = 284.45$ mm. Assuming plano-convex shapes for the first and third element and a biconcave shape for the center element provides a good starting configuration for the computer optimization. Figure 7.4 shows the optimized configuration and the expected blur spots.

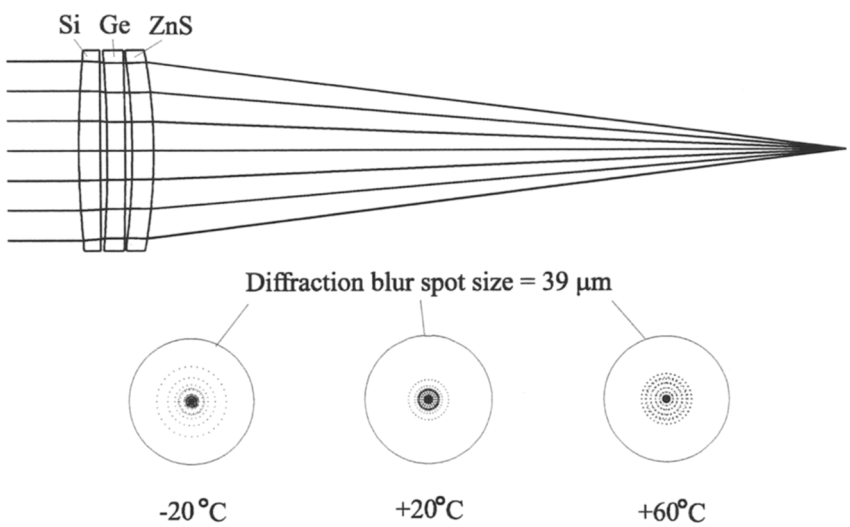


FIG. 7.4 Athermalized achromat (Triplet) for the MWIR region.

7.6.2 The hybrid athermal achromat

For this design, the zinc sulfide element is eliminated.[2] A diffractive phase profile on the rear surface of the germanium lens takes its place. With the reduction from three to two lenses, such an arrangement is cost effective and especially of interest for reducing spherical aberration when the phase profile is superimposed on an aspheric surface, since the asphere and the diffractive structure can be diamond turned at the same time.

As before, the parameters for the first and second element remain as

Element 1: silicon with $v_1 = 236$, and $T_1 = 60\times10^{-6}$ (see Table 7.1)
Element 2: germanium with $v_2 = 103$, and $T_2 = 126\times10^{-6}$.

For the diffractive surface,

$$v_3 = \lambda_M / (\lambda_S - \lambda_L) = 4/(3-5) = -2, \quad T_3 = 2\alpha_L = 6.2\times10^{-6}.$$

We remember that the thermal coefficient of expansion for the aluminum housing is $\alpha_H = 24\times10^{-6}$. Inserting these values into Eqs. (7.9), (7.10), and (7.11) yields focal lengths of $f_1 = 43.810$ mm, $f_2 = -78.299$ mm, and $f_3 = -18{,}332$ mm.

The performance of this athermalized hybrid triplet is shown in Fig. 7.5. The temperature excursion is again ± 40°C.

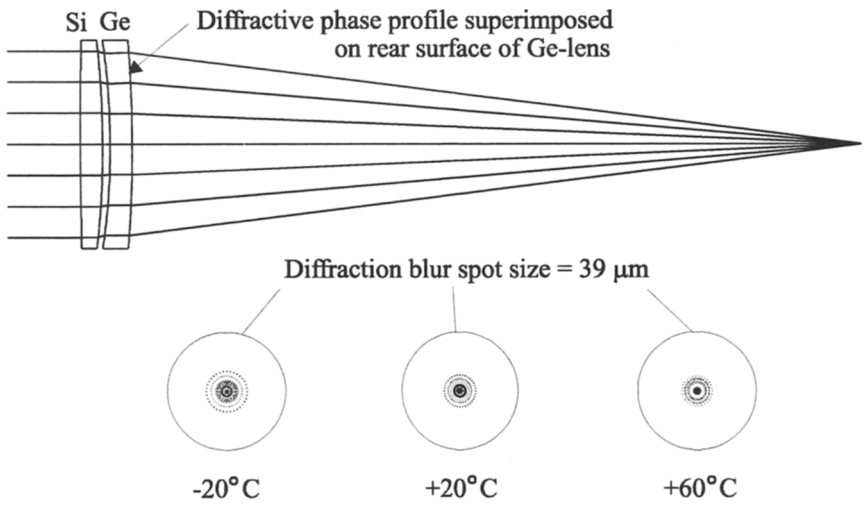


FIG. 7.5 Athermalized achromat (hybrid) for the MWIR region.

7.7 Cold Stop and Cold Shield

7.7.1 Cold stop

In Chapter 2, the cold stop was mentioned as a special stop. As shown in Fig. 7.6, the cold stop is the aperture stop mounted inside the Dewar package. In this arrangement, the detector sees only radiation from the object space. Re-radiation effects from the stop itself are eliminated or at least considerably reduced. One of the most important observations to be made from the Fig. 7.6 is that the cold stop is not only the aperture stop, it is the exit pupil *EP'* as well. The detector is the field stop and the exit window *EW'*, and therefore, with the aperture stop mounted a distance $t_{stop} \equiv S'$ away from the detector, the energy throughput equation [Eq. (2.4)] is already satisfied without any consideration for the imaging optics. We can see that with a round aperture and a square detector,

$$\Gamma = \frac{EP' \times EW'}{S'^2} = \frac{\pi D_{stop}^2 d'^2}{4 t_{stop}^2} = \frac{\pi d'^2}{4(f/\#)^2} . \tag{7.12}$$

The importance of these relations is indicated by the fact that the (*f*/#) of the objective is determined by the detector package geometry and any reduction of the objective's (*f*/#) is meaningless and only wasteful.

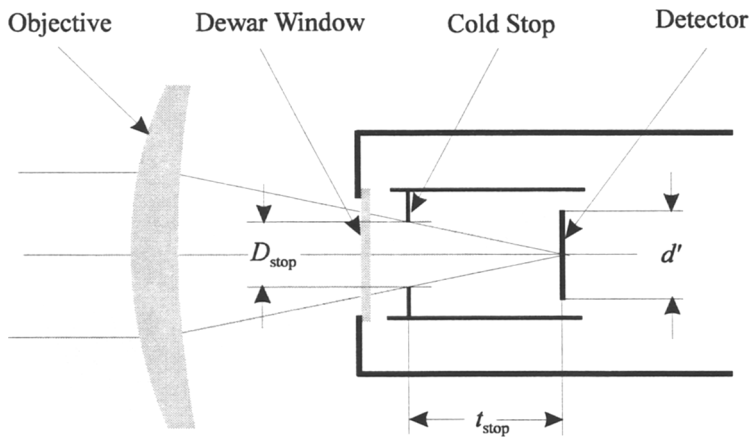


FIG 7.6 The cold stop as system's aperture stop inside the Dewar package.

We also noticed from the simple optical layout that the lens diameter becomes quite large to cover the entire field without vignetting. To solve this problem, reimaging optics can be applied.

7.7.2 Cold shield

If the cold stop is not the aperture stop, it does not control the radiation cone from object space. Depending on the position within the entire optical train, the portion of the

undesired radiation impinging upon the detector varies, and one defines this variation as the cold shield efficiency. A 100% efficiency makes the cold shield a cold stop, because in that case the cold shield has become the aperture stop of the system.[3]

References

1. P. J. Rogers "Athermalization of IR Optical Systems," SPIE Vol. CR 38 (1991) pp. 69-94.
2. R. M. Hudyma "Applications of Diffractive Optical Elements in Infrared Optical Design," SPIE Short Course Notes 1994.
3. R. E. Fischer / B. Tadic-Galeb, *Optical System Design,* McGraw-Hill, Inc. (2000), pp. 221-224.

CHAPTER 8

Optical Coatings

8.1 Introduction

Optical coatings are essential to control the radiation throughput of an optical element. Referring to the basic relation

$$\tau+\rho+\alpha=1, \tag{8.1}$$

where τ = transmittance, ρ = reflectance, and α = absorption, we are primarily interested in the first two contributors. Most of the time, we like to see either a high transmittance for refractive elements (lenses) or a high reflectance for reflectors (mirrors). In the applications for beamsplitters, we like to maximize the efficiency between the transmitted and the reflected energy, which means the absorption must be kept low.

Since the subject of optical coatings is complex and extensive, the coverage will be limited to some fundamental behaviors of thin film coatings and to application remarks. Of special interest for the infrared spectrum are antireflection coatings and bandpass filters. Without these, the entire field of IR optics would be very much restricted.

An important breakthrough came in 1935, when Alexander Smakula from Carl Zeiss in Jena, Germany, received his patent for "a method to increase the transmissivity of optical elements, by reducing the index of refraction on the boundary layers of these optical elements."[1] A few years later, in 1939, Walter Geffcken received his patent for transmission-type interference filters.[2]

8.2 Effects at a Single Surface

For radiation at normal incidence, the reflectance for a single surface is expressed by the fundamental Fresnel equation

$$\rho_s=\left(\frac{N-1}{N+1}\right)^2, \tag{8.2}$$

which remains a good approximation for tilt angles of up to 45°. Neglecting absorption, the transmitted portion is then

$$\tau_s=1-\rho_s=\frac{4N}{(N+1)^2}. \tag{8.3}$$

Figure 8.1 shows the incident ray being separated into a reflected part and a refracted fraction that enters the substrate.

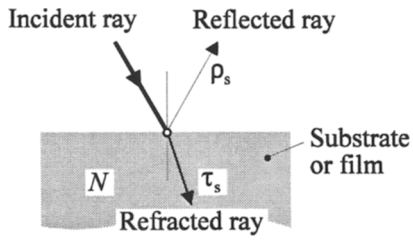


FIG. 8.1 Reflection and transmission on an uncoated single surface.

8.3 Two Plane-Parallel Surfaces

As shown in Fig. 8.2, the effects of reflection and transmission for a substrate with plane-parallel surfaces can be easily determined, even if the internal reflections are included. At each surface boundary there is transmission and reflection. The existing internal reflections can be quite disturbing because these reflections will generate secondary or *ghost* images.

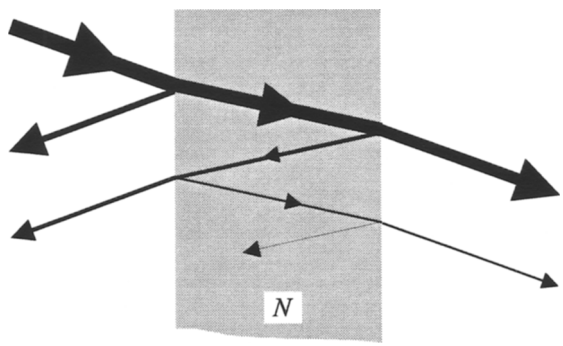


FIG. 8.2 Transmission and reflection at a plane-parallel plate.

By including the internal reflections in the equation for the transmitted fraction of an impinging energy bundle, we find

$$\tau_i = \frac{2N}{\left(N^2+1\right)}. \tag{8.4}$$

If the internal reflections are ignored, the equation changes to

$$\tau_0 = \left[\frac{4N}{(N+1)^2}\right]^2 . \tag{8.5}$$

In comparing these two expressions for different materials, i.e., for different indices of refraction, we notice that for glass with index 1.5, the two values differ by less than 1%. For germanium, with its high index of 4, the numbers contrast by approximately 13%. While this is relatively high, it is more important to notice the absolute transmission differences between a low-index and a high-index material. For example, glass transmits about 92%, but germanium transmits only half of that. That is why an antireflection coating for these high-index materials, when used for IR applications, is absolutely required.

8.4 Antireflection Coatings

The importance of antireflection coatings (AR coatings) can be judged quickly by the number of elements required for a typical zoom lens. Ten elements, or twenty surfaces, can be found easily in such a lens. Assuming all-glass optics with an average index of refraction of 1.5, with a 4% loss per surface, the total transmission would be reduced to approximately 44%. The loss increases dramatically as the index of refraction increases. For germanium, one of the primary materials used in the infrared region, the index of refraction $N \cong 4$. This high index yields a transmission of 41% for only a *single element.*

To correct for a reduction of such magnitude, a method was devised by which the transmitted amount of radiation is increased by reducing the reflection losses. For such an arrangement, the two requirements for a single layer coating, based on interference, are material compatibility for equal amplitude

$$N_F = \sqrt{N_S N_0} , \tag{8.6}$$

and thin film thickness

$$N_F t_F = \frac{1}{4}\lambda , \tag{8.7}$$

where N_F is the index of the thin film, N_S is the substrate's index, and N_0 represents the index of the surrounding medium, which is usually air.

The principle of the interactions at surface boundaries is that there is a half-wave phase change when light travels through a low-index medium and is reflected from a high-index medium. There is no phase change when light travels through a high-index medium and is reflected from a low-index medium. Therefore, as indicated in Fig. 8.3, part of the incident radiation will be reflected at surfaces 1 (beam B) and 2 (beam A), undergoing a half-wave phase change. If the optical thickness of the thin film layer is made one quarter wavelength [Eq. (8.7)], the relative phase shift between the two reflected beams is an

additional half wave which leads to destructive interference (cancellation of amplitude intensities). The two transmitted beams of the incident radiation have also a half-wave phase shift relative to each other due to the quarter-wave film thickness. However, only one beam (beam A in Fig. 8.3) undergoes a phase shift caused by reflection at surface 2. The result is that when these two transmitted beams meet, they will be in phase and their amplitude intensities will be reinforced.

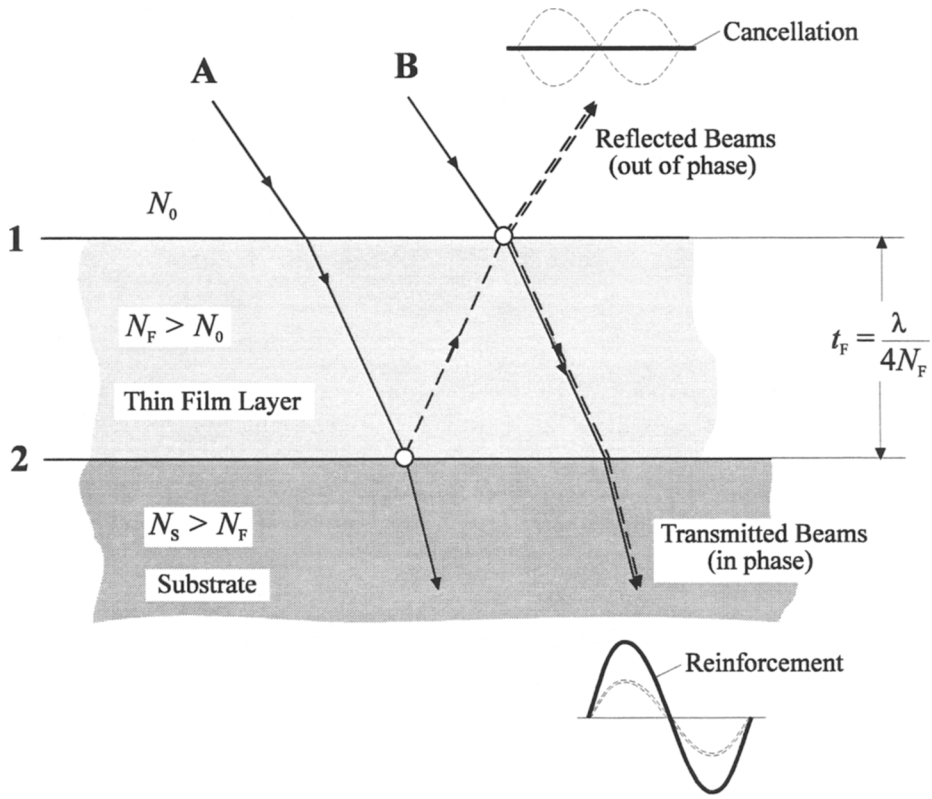


FIG. 8.3 Single-layer antireflection coating.[3] For clarity, the rays have been shown under an angle rather than perpendicular to the thin film and substrate.

An example demonstrates the use of Eqs (8.6) and (8.7). A silicon substrate is to be AR-coated, optimized for λ = 1.7 μm. At that wavelength, the index of refraction for silicon is N_s = 3.47.

According to Eq. (8.6), the index for the film layer should be equal to the square root of the substrate's index N_S if the substrate is in air, for which the index N_0 is 1. The square root of 3.47 is 1.86; this is the index of refraction for silicon monoxide at 1.7 μm. The effect is shown in Fig. 8.4.

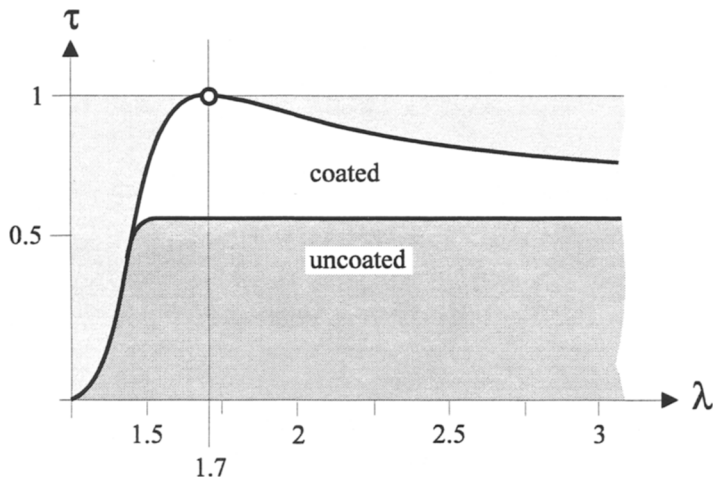


FIG. 8.4 Uncoated and AR-coated silicon surface.

Without the coating, the transmission for this single surface is 69%. Figure 8.4 indicates that the theoretical 100% transmission is achieved only at the specified wavelength. However, the curve does not abruptly plunge down to lower values. Therefore, a certain useful transition band exists. Hudson[4] states that a 10% deviation in optimum index of refraction for the coating [one that meets the square-root relationship of Eq. (8.6)] results in a decrease of only 1% in the efficiency of the coating. This means that if N_F varies from 1.87 to 2.05, for our example, the plate (two surfaces) will still have a total transmittance of at least 98% for the design wavelength of 1.7 μm.

Of course, most applications are not limited to one specific wavelength, but cover a spectral window that can be rather broad. For these cases, a single-layer AR coating is not sufficient. Multilayer coatings have to be used. Their performance depends on the same basic interference principle.

8.5 Reflective Coatings

It was only in the 1930s that the chemical process of "silvering" a surface to create a high-quality optical reflector was replaced by the method of evaporating thin films onto the polished mirror substrate.[5] Today, the most commonly evaporated metal films for reflectors are

- Aluminum
- Silver
- Gold
- Copper
- Rhodium.

Their reflectance variations with wavelengths are summarized in Fig. 8.5.

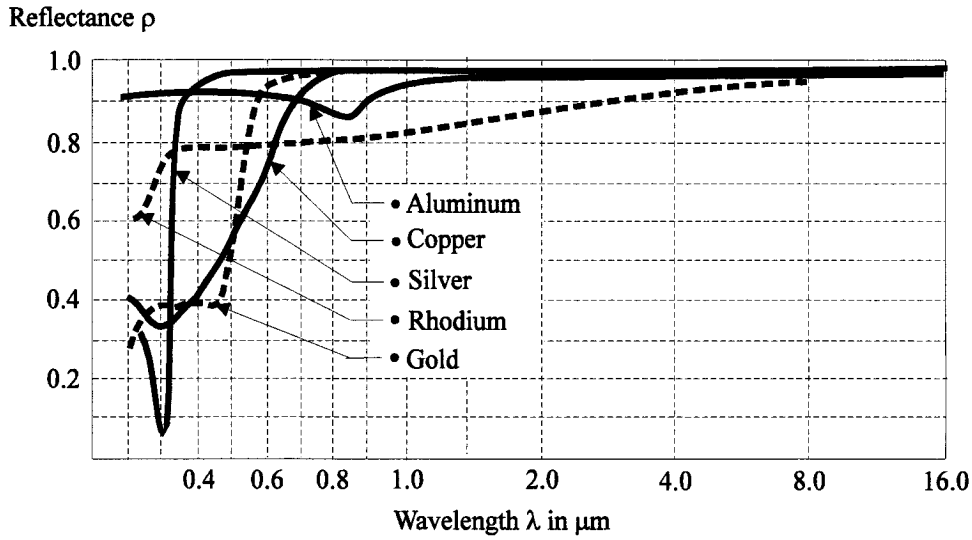


FIG. 8.5 Reflectance of evaporated metal films.

To protect the relatively soft reflective surfaces from the environment, overcoatings are applied. Quite frequently, silicon monoxide or magnesium fluoride is used for that purpose.

Dielectric multilayer coatings are being applied to meet special protective and/or high-reflective requirements. A typical example is the correction for the dip at about 0.85 μm in the aluminum reflectance curve.

8.6 Typical Interference Filters

The principle of interference filters was discussed for the special application for antireflection coatings. IR systems often require transmission control within specific spectral bands. This can be achieved with multilayer interference filters. These filters vary in their complexity according to the performance demand. In some cases, a few thin film layers may be sufficient; in others, 50 or even more may be needed with additional provisions to block radiation to very low levels from being transmitted outside the spectral regions of interest.

Some of the most commonly used filter types are shown in Figs. 8.6–8.9.

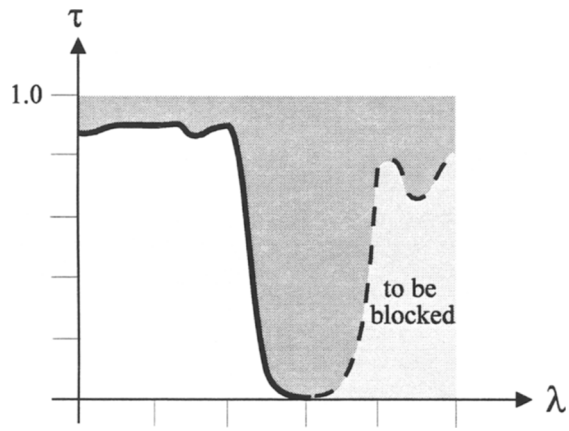


FIG. 8.6 Short-pass filter, (cutoff filter). The branch "to be blocked" requires an additional filter (blocking filter) that absorbs or reflects the radiation otherwise transmitted in that region.

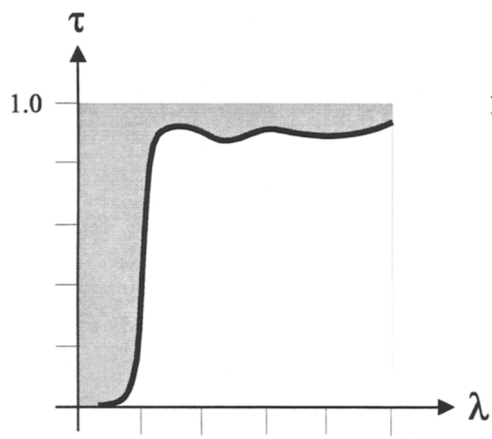


FIG. 8.7 Long-pass filter (cut-on filter).

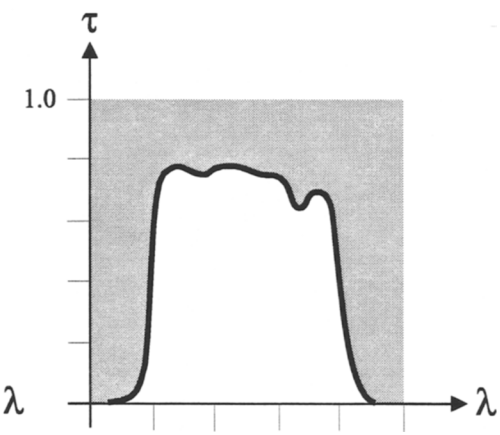


FIG. 8.8 Broad bandpass filter.

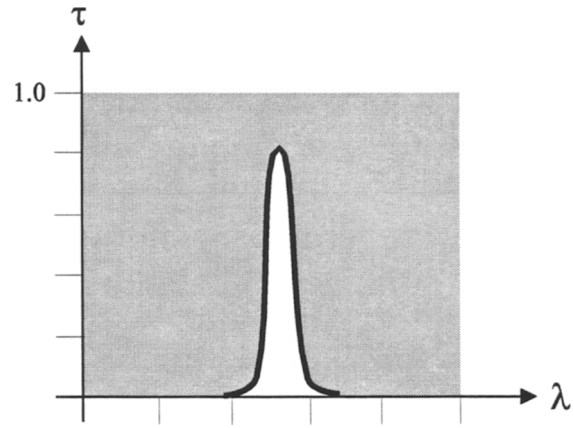


FIG.8.9 Narrow bandpass filter (spike filter).

There are many factors involved in producing these sophisticated filters, and one can sense that neither the design nor the fabrication of optical filters is a trivial matter.[6]

For rugged field spectrometers, where more than just a few selected wavelengths are of interest, linear variable filters (LVF), or circular variable filters (CVF) are employed. Both are based on the behavior of a wedged-film configuration. In the CVF arrangement shown in Fig. 8.10, the wavelength transmitted by the filter varies as a function of the rotation of the filter disk.[7]

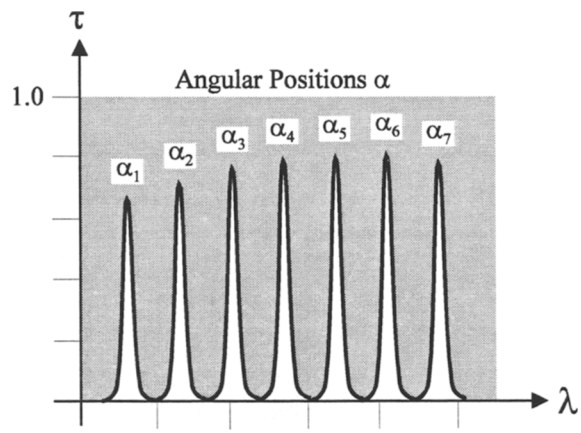


FIG. 8.10 Transmission versus wavelength for CVF.[7]

8.6.1 Angular sensitivity of filters

The performance of a filter changes with the radiation's angle of incidence. For a filter on a plane-parallel substrate tilted by an angle i (see Fig. 8.11), the shift to a different center wavelength is approximated by

$$\lambda_i = \frac{\lambda_0}{N_e}\sqrt{N_e^2 - \sin^2 i}, \tag{8.8}$$

where λ_i is the center wavelength at $i°$, λ_0 is the center wavelength at normal incidence, and N_e is the effective index, derived from indices of thin film layers.

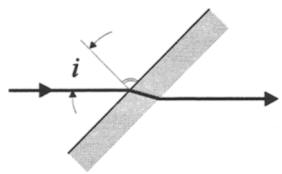


FIG. 8.11 Tilted filter.

From Eq. (8.8) it can be seen that at normal incidence when $i = 0$, the center wavelength λ_i reaches its maximum value. At any other angle there is a shift to *shorter* wavelengths. This feature is sometimes used for spectrally fine-tuning a system.

The angle of incidence at a curved lens surface varies over its entire diameter, resulting in a continuous shift of the center wavelength. While the angular change is most likely never very steep, the fact that a change occurs should not be ignored. This is of course especially important if the lens is not merely coated with an AR film but serves as the substrate for the bandpass filter.

Frequently, a narrow-bandpass filter is inserted in a converging beam near the detector. There are several reasons for that. The physical size of the filter can be kept small, and stray radiation bypassing the filter is greatly reduced. Furthermore, the close proximity to the detector allows for the filter to be included in the cooled Dewar assembly. The variation of angles across the focusing cone has a detuning effect on the filter because of the varying shifts to shorter wavelengths for marginal rays. This is usually considered in the design of the filter, leading to an integrating compromise.

8.6.2 Thermal sensitivity of filters

The index of refraction and the thickness of thin film materials change with temperature. Even though the change of the index N with temperature T (dN/dT) is opposite in sign of the change in thickness t, (dt/dT), the performance of the filter is still affected. Because the first influence is greater than the second, there is a shift toward the *longer* wavelengths as the temperature increases.[8] This is expressed by

$$\lambda_T = \lambda_0 + \Delta T \alpha, \tag{8.9}$$

where λ_T is the center wavelength at temperature T, λ_0 is the center wavelength at temperature T_0, $\Delta T = T - T_0$, and α is the temperature coefficient (μm/°C). (Typical value for $\alpha = 5$ to 12 E-5).

A very useful and extended summary of terms commonly used to describe the spectral characteristics of infrared filters can be found in Ref. 9.

References

1. J. R. Meyer-Arendt, *Introduction to Classical and Modern Optics*, second edition, Prentice-Hall, Inc. (1984), page 274.
2. Ibid., page 271.
3. R. Kingslake, *Optical System Design*, Academic Press (1983), page 106.
4. R. D. Hudson, Jr., *Infrared System Engineering*, John Wiley & Sons (1969), page 219.
5. J. Strong, *Procedures in Applied Optics*, Marcel Dekker, Inc. (1989), page 7.
6. Cerac, Inc., "Coating materials technology: selecting materials and deposition techniques for optical coating," Photonics Design and Application Handbook (1995), page H-81.
7. J. D. Rancourt, *Optical Thin Films Users' Handbook*, MacMillan Company (1987), pages 110-114.
8. Ibid., page 73.
9. Optical Coating Laboratory, Inc., "Infrared filters: glossary, equations, parameters," Photonics Design and Application Handbook (1995), page H-244.

CHAPTER 9

Image Evaluations

9.1 Introduction

With the knowledge of how aberrations and diffraction influence the performance of an optical system, we shall now discuss ways to measure and evaluate image quality. For the infrared, the blur spot size and the radial energy distribution in the image plane are two quantities of great interest. They indicate the minimum detector element size required to collect a certain amount of energy. From this information, one can derive the MTF, a measure of contrast versus resolution.

As with previously discussed subjects, this chapter encompasses a very special and large field. However, the presented approximations will help to quickly obtain some valuable benchmarks for a given situation.

9.2 Blur Spot Measurements

It was pointed out in Chapter 3 that due to diffraction and uncorrected aberrations, the image of an object point is never a point but a diffused disk, usually referred to as the blur circle or blur spot. To confirm the predicted size of a blur spot, several measurement methods are used.

9.2.1 Circular mask

One way of measuring the size of an image blur is to place the lens under test in the exiting beam of a collimator and insert at the image location of the lens a circular mask just large enough to encompass the percentage of the energy the system design asks for. The 100% level is established by recording the signal without any mask. Of course, the detector has to be large enough to accept all the energy transmitted through the lens under test. Next, a mask is inserted, properly sized to the specifications of the desired blur spot size. The signal will indicate if the energy passing through the mask meets the expected percentage level. The principle of such a setup is shown in Fig. 9.1. The use of a field lens provides the necessary clearance for the pinhole aperture insertion.

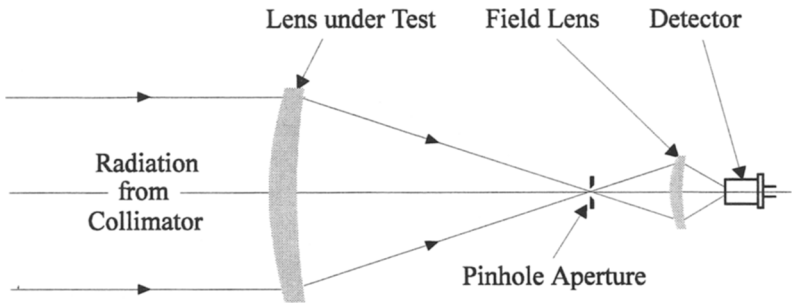


FIG. 9.1 Setup for blur spot measurement with circular masks.

9.2.2 Slit

Substituting a very narrow slit in place of the pinhole and moving it across the image blur is another way of measuring the total blur spot size. To obtain an accurate result, the width of the slit should be no wider than about 1/20 of the blur spot size to be measured. The scanning across the blur spot is done in two directions, orthogonal to each other, if the shape of the blur is not circular.

9.2.3 Knife edge

A third way of measuring the blur spot size is to traverse a sharp edge (knife edge) across the blur. The total spot size and the optimum image position for the minimum blur spot size are found when the travel of the knife edge required to raise the signal from zero to 100% is a minimum. This is indicated in Fig. 9.2.

The question arising from Fig. 9.2 is whether one can simply subtract 10% from both ends of the scan to find the 80% blur spot size. Figure 9.3 shows that this is not the case. Based on the reasonable assumption that the blur spot envelope resembles a Gaussian curve, a correction factor can be applied.

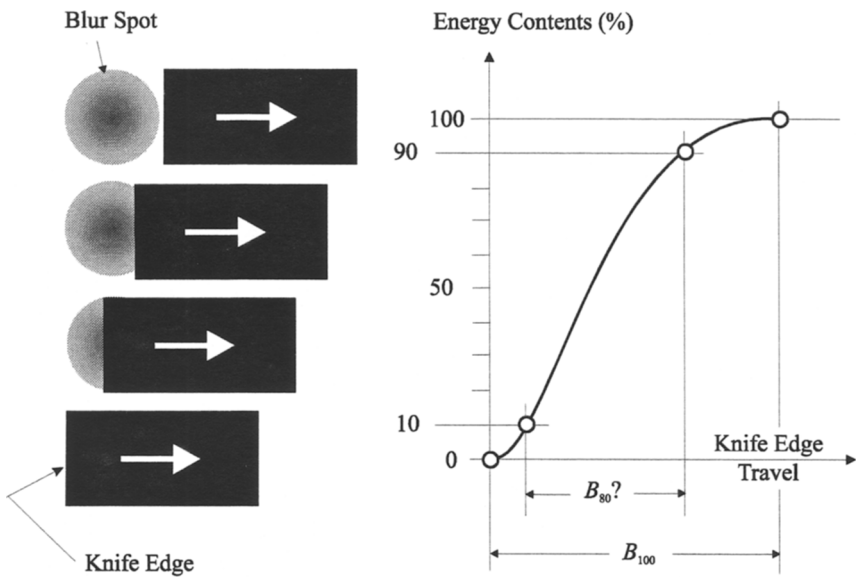


FIG. 9.2 Knife edge scan across the blur spot.

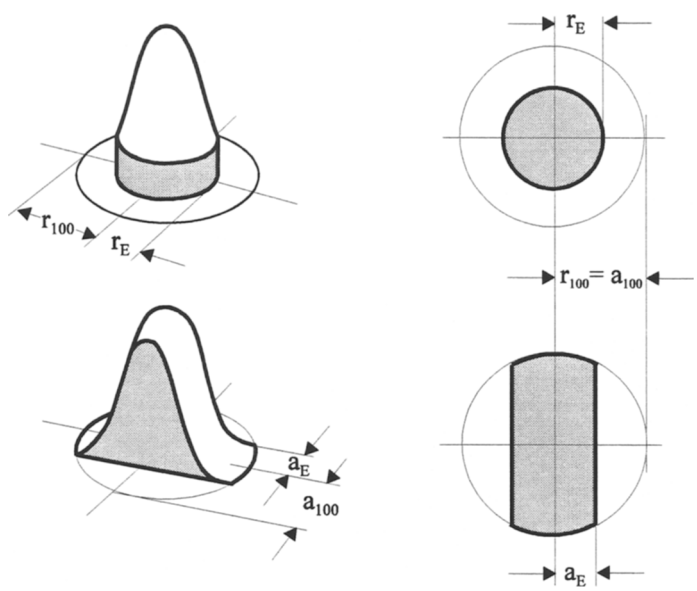


FIG. 9.3 Comparison between measurements with circular masks and knife edge scan.

The relation between the distance a and the radius r for the same amount of energy E contained in the exposed portion of the blur spot is approximated by

$$r \cong 1.25a\,E^{-0.477}. \tag{9.1}$$

For example, for E = 80%, $r = 1.25a\ 0.80^{-0.477} = 1.39a$. In other words, if the dimension a obtained with the knife measurement was 0.2 mm, the radius of the circular blur spot would be $0.2 \times 1.39 = 0.28$ mm. The main point to remember is that one has to be careful with the interpretation of data obtained by one measuring method when applied to another.

9.3 Energy Distribution

To properly assess the energy distribution of a blur spot, measurements are required that indicate the change of the encircled energy as a function of the blur spot radius. The setup shown in Fig. 9.1 is suitable for such a measurement. A number of well-centered circular masks of increasing size are inserted successively at the image location and the energy levels are recorded from 0 to 100%. The resultant plot looks typically like the one shown in Fig. 9.4.

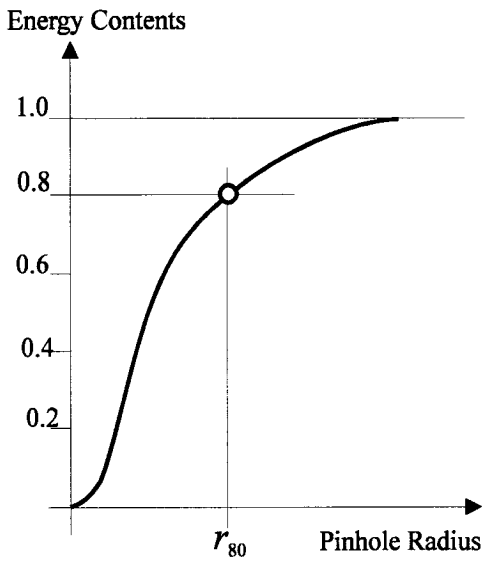


FIG. 9.4 Radial energy distribution.

9.4 Modulation Transfer Function

Removing the subjectivity from evaluating lenses by just visually judging their performance was achieved with the introduction of the modulation transfer function. This occurred about a half century ago.[1]

While the subject of MTF is very complex, there are a number of approximations that are especially helpful in the layout stages to assess the performance expectations of a

lens or lens system. Before we discuss these, let us look at some basic principles related to the MTF.

9.4.1 Overview

The image quality of an object is a measure of the quality of a lens. With the MTF the image quality is measured as contrast against spatial frequency.

Contrast or modulation is expressed by

$$M = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}. \tag{9.2}$$

For the object, $I_{O\,max}$ and $I_{O\,min}$ are the maximum and minimum intensities of the radiant emittance and for the object. $I_{I\,max}$ and $I_{I\,min}$ relate to the irradiance of the image. This is identified for a typical bar pattern in Fig. 9.5.

The modulation transfer function [MTF(ν)] is the ratio of the modulation of the image [M_i (ν)] and the modulation of the object [M_o]. ν is the spatial frequency in line pairs (one dark and one light line) per unit length, usually millimeter. (Other dimensions, such as cycles per radian, are used as well). The number of line pairs is increased until the contrast in the image is too low to be detected and the imaged pattern can no longer be resolved. At that "limiting resolution" point, the spatial frequency is called the cutoff frequency ν_0. For an aberration-free system,

$$\nu_0 = \frac{1}{\lambda\left(f/\#\right)}. \tag{9.3}$$

Fig. 9.6 shows how the modulation reduces as the spatial frequency increases.

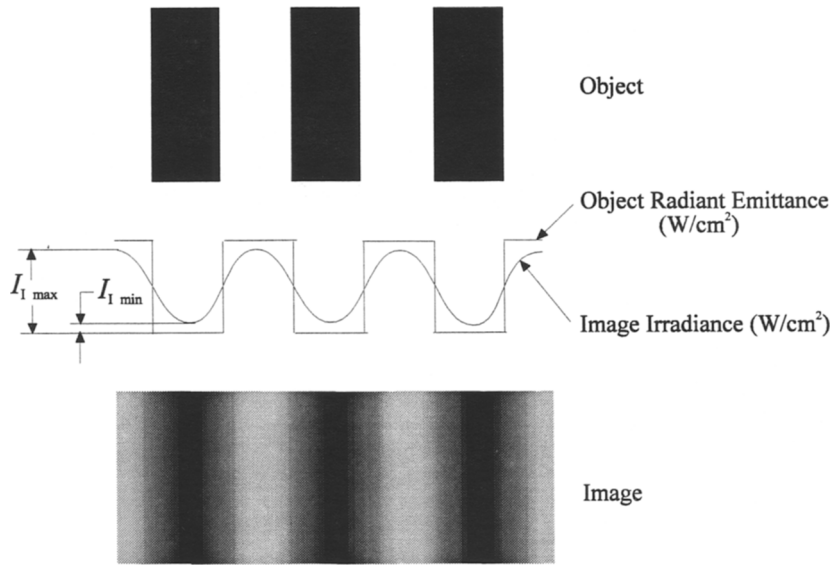


FIG. 9.5 Imagery of a bar pattern with $I_{O\,max} = 1$ and $I_{O\,min} = 0$ ($M_O = 1$).

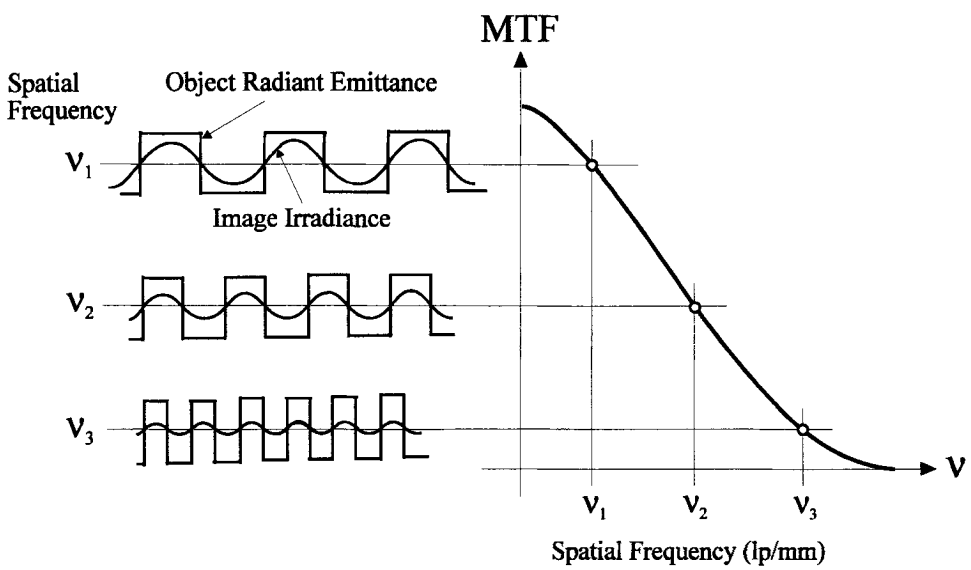


FIG. 9.6 The modulation transfer function.

MTF analyzers are being offered in a variety of configurations. Whether slits, knife edges, sinusoidal, or bar pattern targets are being used, the basic principle is always the same:

> ***The modulation of the image is compared with the modulation of the object at various spatial frequencies.***

9.4.2 Contrast and resolving power

As mentioned earlier, contrast is the relationship between two shades of gray which, at the limit, becomes black and white. Resolution is the limit of discernibleness of two fine structures, such as the imaged lines of the MTF analyzer's bar pattern target. With these two definitions, it can be understood that neither contrast nor resolution alone describe the full performance of an optical system. This is indicated in Fig. 9.7 where the modulation transfer functions of two lenses are compared. Lens A has higher modulation (contrast) at lower frequencies than Lens B. On the other hand, Lens B has a higher cutoff frequency (resolution limit) than Lens A. The question is not which lens is better, it is which lens is more suitable for the intended application. The curves provide the answer.

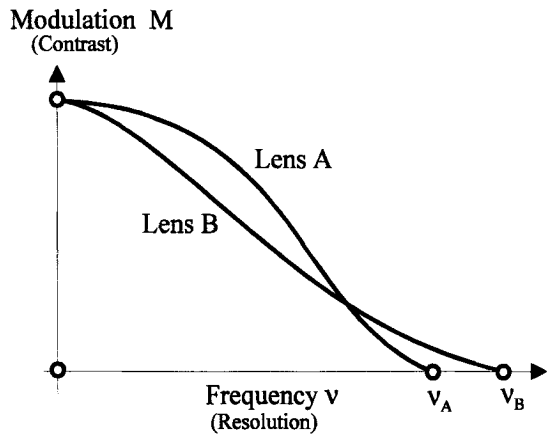


FIG. 9.7 Contrast and resolving power, two general cases.

Erich Heynacher and Fritz Köber, both with Carl Zeiss, Inc. have published an interesting paper on this subject of contrast and resolution.[2] With their kind permission, we show reproductions in Fig. 9.8 of four photographs of the same subject (Emperor Henry II in a stained glass window at the Strassburg cathedral) taken with four different lenses. While there is of course a degradation in the quality of the reproductions presented here, the relation between contrast and resolution is well enough preserved to demonstrate the principle indicated in Fig. 9.7. As the authors of that paper suggest, a very educational experiment is to observe the pictures from a larger viewing distance. As the distance is increased, the resolution limit of the eye becomes the limiting factor and confirms again that evaluating or judging the image quality of an optical system is not just a matter of the resolving power of the lens. All of this can be applied to IR imaging systems where the eye is replaced by a detector.

1 • Poor contrast
• Reasonable resolution

2 • Better contrast
• Worse resolution

3 • Better contrast than 1, but worse than 2
Best resolution

4 • Best contrast
• Same resolution as 1

FIG. 9.8 Contrast and resolving power.

The associated MTF curves shown in Fig. 9.9 confirm clearly the characteristics of the four lenses used to take the pictures shown in Fig. 9.8.

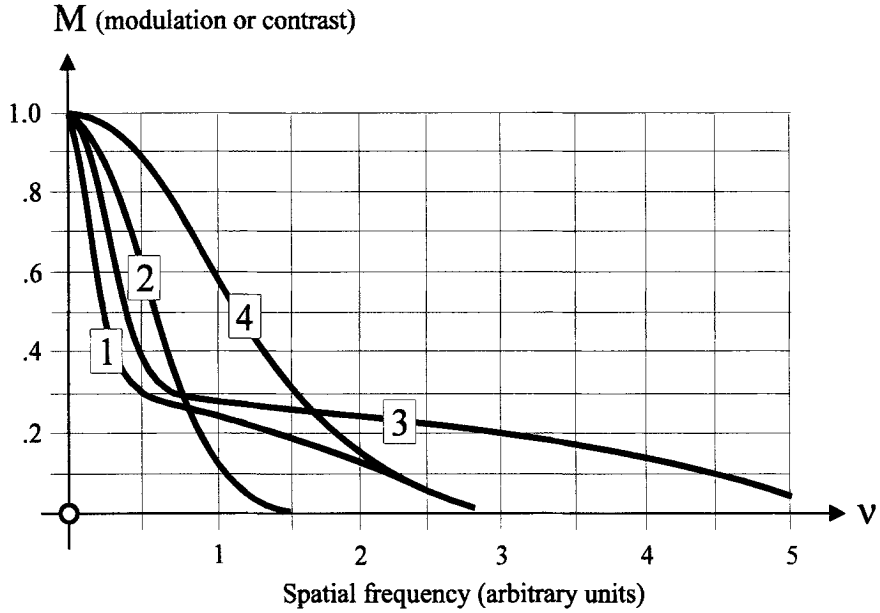


FIG. 9.9 Contrast and resolving power, the transfer functions.

That resolution is only part of the information for judging image quality is evident from the shapes of curves 1 and 4. Both show the same cutoff frequency, i.e., the same limiting resolution, but vary strongly with regard to contrast. This confirms the remarks made in Fig. 9.8. Another observation substantiates the visual renditions of samples 1 and 2. The limiting resolution of sample 2 is about half the one of sample 1; however, its contrast is much higher at lower frequencies. All this indicates that much can be predicted and confirmed with modulation transfer functions.

A typical MTF analyzer is shown in Fig. 9.10. Such a system can be equipped with a number of convenient attachments for evaluating a lens at different conjugates and off-axis positions. Polychromatic tests can be conducted and defocusing effects can be analyzed.

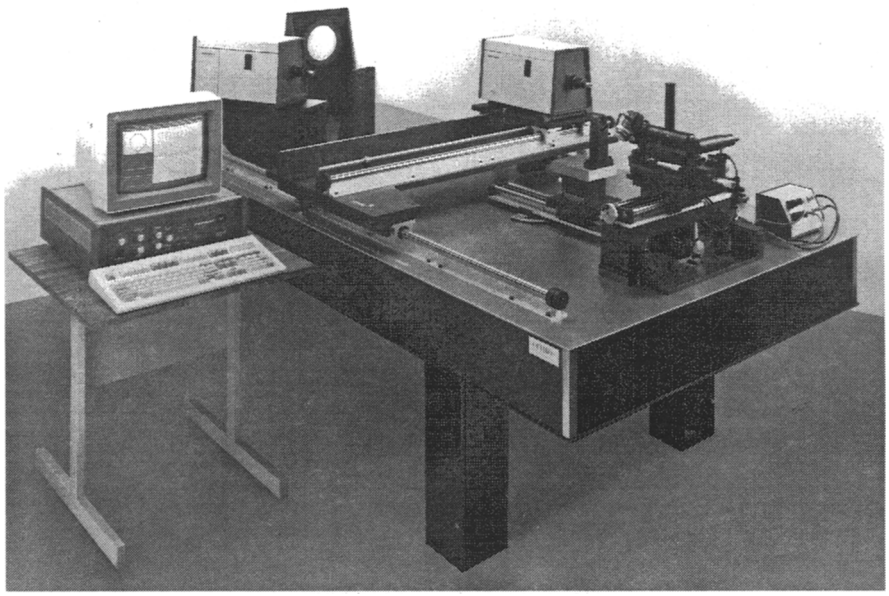


FIG 9.10 Commercial MTF analyzer
(Reprinted with permission of Optikos Corporation)

9.4.3 Diffraction MTF

For a diffraction-limited optical system, which is one with negligible aberration effects, the MTF is stated by

$$MTF_{\text{diffr}}(\nu) = \frac{2}{\pi}\left[\arccos\left(\frac{\nu}{\nu_0}\right) - \left(\frac{\nu}{\nu_0}\right)\sqrt{1-\left(\frac{\nu}{\nu_0}\right)^2}\right]. \qquad (9.4)$$

This equations refers to an optical system with a circular aperture. As stated before, ν_0 is the cutoff frequency of the system, expressed by

$$\nu_0 \cong \frac{1}{\lambda(F/\#)}. \qquad (9.5)$$

This simple relation provides us with another "convenient to remember" fact. Selecting 0.5 μm as the wavelength for the visible spectrum, 4 μm for the MWIR, and 10 μm for the LWIR leads to the following limits:

$$\nu_{0\,\text{VIS}} = \frac{2000}{(f/\#)},\ \nu_{0\,\text{MWIR}} = \frac{250}{(f/\#)},\ \text{and}\ \nu_{0\,\text{LWIR}} = \frac{100}{(f/\#)}. \qquad (9.6)$$

These relations refer to line pairs per millimeter. It is good to remember that even a diffraction-limited ($f/1$) system in the LWIR cannot resolve more than 100 lp/mm.

Within the boundaries of $0 < (\nu/\nu_0) < 0.6$, a good linear approximation for the diffraction MTF is given by

$$MTF_{\text{diffr}} \cong 1 - 1.218\left(\frac{\nu}{\nu_0}\right). \tag{9.7}$$

Also, using Eq. (6.5),

$$MTF_{\text{diffr}} \cong 1 - 1.218\lambda\,(f/\#)\nu. \tag{9.8}$$

9.4.4 Geometric MTF

The calculation of the geometric MTF assumes a perfect optical system and judges the performance of a lens as a function of defocusing. Diffraction effects are ignored. This assumption leads to

$$MTF_{\text{geom}}(\nu) = \frac{2J_1(\pi\nu B)}{\pi\nu B} \cong \frac{2J_1\left[\pi\nu\delta/(f\quad/\#)\right]}{\pi\nu\delta/(f\quad/\#)}, \tag{9.9}$$

with J_1 the first-order Bessel function, B the linear blur spot size, and ν the spatial frequency. δ is the defocusing necessary to yield blur spot B, i.e.,

$$\delta = B(f/\#). \tag{9.10}$$

Expression (6.8) is not too reliable for small amounts of defocusing.[3] A better result is obtained by using an approximation that includes a defocusing multiplying factor. J. H. Haines and J. P. King,[4] as well as R. R. Shannon[5] have discussed the approximation adopted here.

$$MTF_{\text{appr}} = MTF_{\text{diffr}} \times DMF. \tag{9.11}$$

The defocusing multiplying factor

$$DMF = \frac{2J_1\left\{\pi\nu B\left[1 - (f/\#)\lambda\nu\right]\right\}}{\pi\nu B\left[1 - (f/\#)\lambda\nu\right]}. \tag{9.12}$$

9.4.5 Numerical example

Given: Lens with $f = 100$ mm, $f/\# = 2$, wavelength $\lambda = 4$ μm.

Determine: Diffraction blur spot size, cutoff frequency,
modulation at 30 lp/mm for diffraction-limited system
and for system with 0.32 mrad angular blur spot size.
Plot MTF curves, including approximations.

1. Diffraction blur $B_{\text{diffr}} = 2.44 \times 0.004 \times 2 = 0.0195\,\text{mm}.$

2. Cutoff frequency $$\nu_0 = \frac{1}{0.004 \times 2} = 125\,\text{lp/mm}.$$

3. Normalized frequency $$\frac{\nu}{\nu_0} = \frac{30}{125} = 0.24.$$

4. Modulation $$MTF_{\text{diffr }30} = \frac{2}{\pi}\left(\arccos 0.24 - 0.24\sqrt{1-0.24^2}\right) = 0.70.$$

5. Approximation $$MTF_{\text{diffr appr }30} \cong 1 - 1.218 \times 0.24 = 0.71.$$

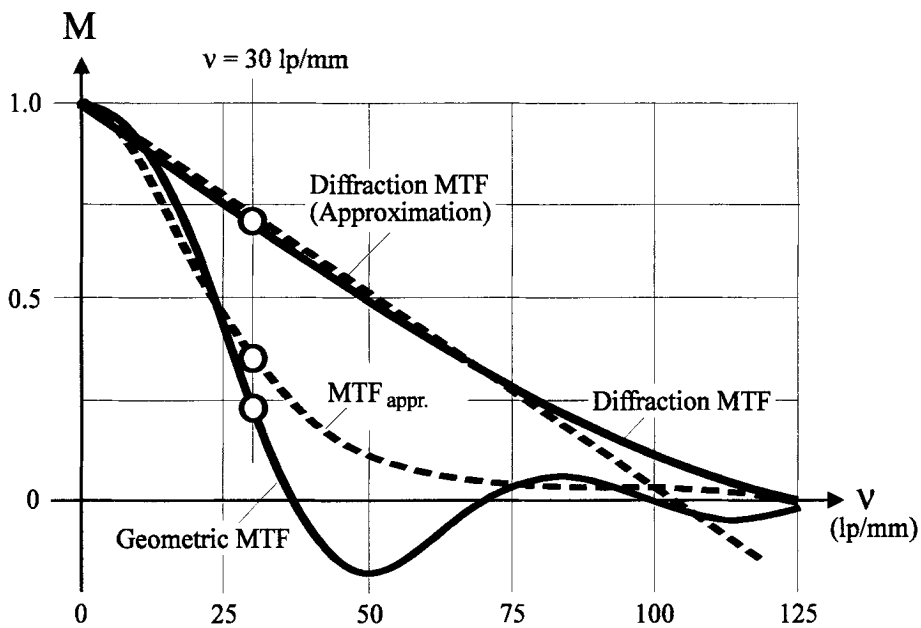
6. The angular blur of 0.32 mrad becomes a linear blur of 0.00032×100 = 0.032 mm in the focal plane and results in

$$MTF_{\text{geom}} = \frac{2J_1(30 \times 0.032\pi)}{30 \times 0.032\pi} \cong 0.22\ .$$

7. Improved approximation, applying correcting method

$$MTF_{\text{appr }30} = 0.70 \times \frac{2J_1[30 \times 0.032\pi(1 - 2 \times 0.004 \times 30)]}{30 \times 0.032\pi(1 - 2 \times 0.004 \times 30)} \cong 0.33\ .$$

8. MTF plots



Remark: The geometric MTF curve shows negative values. This indicates a 180° phase shift in the image. The effect is a dark image where it should be light and vice versa. This is known as spurious resolution.

References

1. C. S. Williams and O. A. Becklund, *Introduction to the Optical Transfer Function*, John Wiley & Sons (1989), page 9.
2. Erich Heinacher and Fritz Köber, "Resolving power and contrast," Zeiss Information No 51, 12th year, pages 29 to 32.
3. W. J. Smith, *Modern Optical Engineering*, McGraw-Hill, Inc. (2000), page 379.
4. J. H. Haines and J. P. King, "Approximate Computation of Physical MTF," SPIE proceedings, vol. 13 (1968), pages 105 to 113.
5. R. R. Shannon, "A useful optical engineering approximation," *Optics and Photonics News*, (April 1994), page 34.

INDEX

Abbe, 49
Aberration,1, 4, 16
Aberration contributions, 69, 86
Aberration correction methods, 53
Achromat, 56, 109
Air-spaced elements, 57
Airy disk, 75
Angular blur spot size, 44
Angular sensitivity, 148
Antireflection coating, 143
Aperture stop, 1, 24
Appendix, 173
Aspheres, *(see General aspheres)*
Astigmatism, 37, 39, 48
Athermal achromat, 135
Athermalization, 133
Atmospheric transmission, 12, 13
Axial chromatic aberration, 37, 41, 49

Back focal length, 62
Background, 4
Ball lens, 93
Beam expanders, 59, 73
Best fit sphere, 98
Binary optics, 107, 117
Blackbody, 5
Blur spot size, 151, 178

Cassegrain configuration, 67, 68
Chromatic aberration, *(see Axial chromatic aberration)*
Circle of least confusion, 40, 47, 50
Circular masks, 151
Classical two-mirror configurations, 67
Cold shield, 138
Cold stop, 138
Coma, 37, 38, 39
Conic sections, 97
Contrast, 157, 159
Curvature, 43, 47

D*, 3, 15, 16
Design examples, 121
Detector configurations, 16
Diamond-turning, 165
Differential changes, 29
Diffraction efficiency, 114, 115
Diffraction limit, 74
Diffraction MTF, 160
Diffractive optics,107
Dispersion, 49
Distortion, 37, 41
Domes, 91

Effective focal length, 62
Emissivity, 4, 7
Encircled energy, 154
Energy transfer, 28
Entrance pupil, 1, 2
Entrance window,1, 2
Etendue, 27
Euler, 45
Exit pupil, 1, 2
Exit window, 1, 2,

f-number, 2
Field curvature, 37, 40
Field flattener, 72, 73
Field lens, 34
Field of view, 16, 35
Focal length, 2

Gauss, 45
General asphere, 103
Geometric MTF, 161
Germanium, 43, 46, 61
Gradient index lens, 95
Gregory configuration, 71

Humidity, 11
Hybrid achromat, 109

Image formation, 21
Immersion lenses, 30
Infrared regions, 9, 49
Instantaneous field of view, 128
Interference filters, 146
Internal reflections, 143
Irradiance, 155

Kirchhoff, 7
Knife edge, 152

Lateral color, 42
Light pipes, 33
Longitudinal spherical aberration, 38

Machining time, 171
Magnification, 23
Mangin, 65
Marginal ray, 1, 59
Minimum blur, 50
Modulation transfer function, 156

Net curvature, 42
Noise equivalent bandwidth, 15
Noise equivalent power, 3

Optical gain, 30, 118
Optical materials for the IR, 178
Optical micrometer, 85
Optical path length, 78
Optical path length difference, 78
Optical surface testing, 169

Paraxial ray tracing, 173
Performance equation, 4
Petzval, 40, 41, 57, 125
Photon detectors, 14
Planck, 5, 6
Plane-parallel plate, 83, 86
Power equations, 175
Precipitable water, 10
Prism, 20, 90
Pupils, 24

Radiance, 2, 3
Radiant emittance, 3, 4
Rayleigh, 75
Reflective beam expanders, 73
Reflective coatings, 145
Reflective optics, 62
Refractive beam expanders, 59
Refractive index, 42
Relative humidity, 11, 12
Resolution, 75

Scattering, 168
Schwarzschild, 68, 72
Seidel, 42
Signal-to-noise ratio, 2, 4
Silicon, 46, 61
Slit, 152
Snell's law, 19, 20
Spatial frequency, 155
Specific detectivity, 15
Spectral radiant emittance, 5
Spherical aberration, 37
Spherical mirror, 63
Spike filters, 148
Staring arrays, 34
Stops, 24
Surface finish. 166

Telephoto, 62
Thermal detectors, 14
Thermal effects, 131
Thermal sensitivity, 149
Thin lens, 21, 43
Three-mirror configurations, 105
Throughput, 26
Transmission, 143
Transverse spherical aberration, 44

Wave aberrations, 77
Wavelength, 5, 6
Wien's displacement law, 7

Virtual image, 23, 24

Zinc selenide, 46, 61

About the Author

Max J. Riedl was born in Kempten, Germany, where he was educated at the Mathematisch Mechanisches Institut. He also graduated from the Akademie für Angewandte Technik, where he studied precision mechanics and optics. He has worked in the field of mathematical and optical instruments for many years and was awarded numerous patents for mechanical, optical and electro-optical devices. He has held technical as well as business leadership positions during the 45 years he lived in the United States. Riedl has published and presented many papers on fundamental lens design as well as on replicating optical elements and diamond turning diffractive optical components for applications in the infrared spectrum. He teaches short courses on these subjects on a regular basis and is a Fellow of SPIE. Presently, he resides in Bavaria, Germany.